Title
Three Essays about Higher Education Programs and Policies to Support Student Success

Author
Solanki, Sabrina

Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Three Essays about Higher Education Programs and Policies to Support Student Success

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Educational Policy & Social Context

by

Sabrina Mahendra Solanki

Dissertation Committee:
Assistant Professor Di Xu, Chair
Distinguished Professor Greg Duncan, Co-Chair
Distinguished Professor George Farkas
Assistant Professor Rachel Baker

2019

DEDICATION

I dedicate this to my family and friends.

To my mother and father. You both made a courageous journey to the United States decades ago. Because of your courage, I have lived a life full of opportunity and have had experiences that have brought me so much joy. And, because of your unrelenting support, I have been able to achieve all of my dreams and look forward to my future with great optimism. In everything that I do, I think about making you both proud. I appreciate your courage and support and all that you have done and sacrificed so that I can live this great life. This Ph.D. is dedicated to the both of you.

To my sister, Sonia. Whenever I have doubted myself, you have been there to remind me that I have what it takes to keep going. Thank you for always believing in me! And to my brother-in-law, Sean. I appreciate your genuine curiosity and openness to learn. Thank you for always being interested in my work and for being available to listen.

To Jimmy. You are a great coach and life partner. More than anyone else, you have seen how challenging this journey has been for me. But your constant support and dedication to my career helped me make it to the finish line. I am forever grateful to you for making my dreams come true.

To my friends who have been by my side since childhood. I appreciate all of you and am grateful for your friendship, support, and encouragement!

Dimple, I appreciate the endless hours you have spent discussing my research, my goals, and my trials and tribulations. Thank you for always being around to listen and to advise.

Tiffany, I know that I've made it this far because of the good energy you've sent my way and because you've kept me in your thoughts. Thank you!

Dimitra, thank you for helping me keep things in perspective and maintain a sense of rationality throughout this journey that has often taken me down many deep rabbit holes.

TABLE OF CONTENTS

	Page
LIST OF FIGURES	iv
LIST OF TABLES	v
ACKNOWLEDGMENTS	vii
CURRICULUM VITAE	viii
ABSTRACT OF THE DISSERTATION	xiii
INTRODUCTION	1
CHAPTER 1: Charting College Success: A Meta-Analytic Study of Social-Psychological Interventions	8
CHAPTER 2: Learning Communities in STEM Education: Evaluating the Impact of the EASE Program on STEM Success	59
CHAPTER 3: Tenure-Track Appointment for Teaching-Oriented Faculty: The Impact of Teaching and Research Faculty on Student Outcomes	110
CHAPTER 4: Summary and Conclusions	165

LIST OF FIGURES

		Page
Figure 1.1	PRISMA Flow Chart	56
Figure 1.2	Funnel Plot	57
Figure 1.3	Histogram of Effect Size Estimates	58
Figure 2.1	Probability of EASE Participation by SAT Math Test Score	104
Figure 2.2	RD Validity Check : SAT Math Score Distribution by Pretreatment Individual Characteristics	104
Figure 2.3	RD Validity Check: Density of SAT Math Score Distribution Around the Cutoff Score	105
Figure 2.4	Density Test	105
Figure 2.5	Trends in EASE Participation and Cohort Composition Before And After Implementation of the EASE Program	106
Figure 2.6	Outcome Measures by SAT Math Score	107
Figure 3.1	Faculty Distribution by Field of Study	159

LIST OF TABLES

		Page
Table 1.1	Mean Effect Size Estimates	50
Table 1.2	Institution Selectivity Categories	52
Table 1.3	List of Studies Included in Meta-Analysis	53
Table 1.4	Review of Literature & Comparison of ESEs	55
Table 2.1	Summary Statistics: 2016 & 2017 Biological Sciences Majors	96
Table 2.2	Covariate Balance Check	97
Table 2.3	RD Results: Social-psychological Measures of the Student Experience	98
Table 2.4	RD Results: First-Year Biology Courses	99
Table 2.5	RD Results: End of Year 1 Academic Outcomes	100
Table 2.6	RD Results: Outcome Measures—Differences by Cohort	101
Table 2.7	RD Results: End of Year 2 Academic Outcomes	102
Table 2.8	RD Results: Treatment Effect Heterogeneity	103
Table 2.9	DID Results: Treatment Effect Heterogeneity	104
Table 3.1	Student Descriptive Statistics	151
Table 3.2	Instructor Descriptive Statistics	152
Table 3.3	Summary Statistics for Outcome Measures by Faculty Type	153
Table 3.4	Probability of Taking an Introductory Course with Different Types Instructors	154
Table 3.5	Contemporaneous Course Outcomes	156
Table 3.6	Impact of Different Types of Instructors on Course Performance: Grade Distribution	157
Table 3.7	Subsequent Outcomes	158

LIST OF APPENDIX TABLES

		Page
Appendix Table 2.1	End-of Term Survey by Construct	108
Appendix Table 3.1	Faculty Titles by Faculty Type	160
Appendix Table 3.2	Faculty Distribution by Field of Study	161
Appendix Table 3.3	Summary Statistics: Analytical Sample and Restricted Sample	162
Appendix Table 3.4	Impact of Different Types of Instructors on Total Credits	163
Appendix Table 3.5	Relevant Literature	164

ACKNOWLEDGMENTS

There are far too many nice things to say about my mentors and colleagues at UCI. Di Xu, my doctoral advisor, has been an incredible source of support. I appreciate the high standards she has set for me. Her constant encouragement and support helped me reach each milestone, including this one. And she has always been someone that I can count on. More than anyone else, she has shaped the way I think about my work and the way I approach it. She has been an amazing mentor and teacher. I am glad the completion of this document marks not the end of our work together but the beginning of a new stage of collaboration.

I would also like to thank the other three members of my dissertation committee: Rachel Baker, Greg Duncan, and George Farkas. Their commitment, guidance, and feedback throughout my journey at UCI helped me develop my identity as a researcher. I have learned a great deal from them and consider them all to be excellent mentors and teachers.

I feel fortunate to have spent the past five years with amazing peers at UCI. Without a doubt, they have been crucial to my development and growth as a scholar. They have also been a necessary support system to make it through a Ph.D. program. I appreciate the time that I have spent with them in classes, collaborating on projects, learning about their work, and going through each milestone together. I have grown smarter as a result of my many interactions with them.

I would also like to thank Phoebe Camilletti. I have learned so much from Phoebe's feedback on my written work; she has helped me think deeply about the words that I put down on paper. And has been integral to my growth as a writer.

Lastly, the research described in this dissertation has been funded by the NAEd/Spencer Foundation. I thank them for their generous support.

CURRICULUM VITAE

Sabrina M. Solanki

ssolanki@uci.edu
sabrinasolanki.com

EDUCATION

Ph.D., Educational Policy & Social Context, University of California, Irvine, CA, 2019

M.A., Public Policy, University of California, Irvine, CA, 2014

M.A., Teaching with a Single Subject Credential in Social Science, specialization in Economics, University of California, Irvine, CA, 2007

B.A., Economics, University of Caifornia, Los Angeles, CA, 2001

FELLOWSHIPS

2018 National Academy of Education/Spencer Dissertation Fellowship ($27,500)

PUBLICATIONS

Solanki, Sabrina M. and Di Xu. (2018). Looking beyond academic performance: The influence of instructor gender on student motivation in STEM fields. *American Educational Research Journal, 55*(4), 801-835.

Xu, Di, **Sabrina M. Solanki**, Peter McPartlan, and Brian Sato. (2018). EASE'ing students into college: The impact of multidimensional support for underprepared students. *Educational Researcher, 47*(7), 435-450.

Solanki, Sabrina M., Peter McPartlan, Di Xu, and Brian Sato. (2019). The impact of a STEM learning community: Does it work and which students benefit the most? *PLoS ONE, 14*(3):e0213827.

Penner, Emily K., Jane Rochmes, Jing Liu, **Sabrina M. Solanki,** and Susanna Loeb. Equity-oriented applicants: What do prospective teachers say about the achievement gap and does it make them more attractive candidates? (available online). *RSF: The Russell Sage Foundation Journal of the Social Sciences (Special issue on using administrative data for science and policy)*.

Xu, Di, **Sabrina M. Solanki**, and Ashley Harlow. Transferring from a community college or initiating in a four-year college directly: Interpreting the effects in the presence of heterogeneity. (In press). *Research in Higher Education*.

MANUSCRIPTS UNDER REVIEW

Xu, Di and **Sabrina M. Solanki**. Tenure-track appointment for teaching-oriented faculty? The impact of teaching and research faculty on student outcomes. Revise and resubmit, *Educational Evaluation and Policy Analysis.*

MANUSCRIPTS IN PREPARATION

Solanki, Sabrina M., Masha R. Jones, and Hansol Lee. Charting college success: A meta-analytic study of social-psychological interventions.

Solanki, Sabrina M. Learning communities in STEM education: Evaluating the impact of the Enhanced Academic Success Experience (EASE) initiative on student success.

Penner, Emily, **Sabrina M. Solanki,** and Xuehan Zhou**.** First impressions: Can initial district screenings of applicants predict teacher outcomes on the job?

Baker, Rachel, **Sabrina M. Solanki,** and Di Xu**.** The effects of free classes over the summer: A fee incentive program to increase summer session enrollment.

Xu, Di, John Fink, and **Sabrina M. Solanki.** The geography of racial gaps in advanced placement participation and dual-credit education.

Baker, Rachel, **Sabrina M. Solanki,** and Connie Kang**.** Racial segregation in California higher education: Examining student population trends between and within sectors.

INVITED PRESENTATIONS

Do Learning Communities Work in STEM Education: Evaluating the Impact of the Enhanced Academic Success Experience Initiative (EASE) on Student Success. ExpandED: Broadening the Understanding of Today's Educational Issues. April 2018. University of California, Santa Barbara.

CONFERENCE PRESENTATIONS

The Effects of Free Classes Over the Summer: A Fee Incentive Program to Increase Summer Session Enrollment. Association for Education Finance and Policy (AEFP) Annual Meeting. March 2019. Kansas City.

First Impressions: Can Initial District Screenings of Applicants Predict Teacher Outcomes on the Job? Association for Education Finance and Policy (AEFP) Annual Meeting. March 2019. Kansas City.

Tenure-track for Teaching-Oriented Faculty: Faculty Characteristics, Growth, and

Impacts on Undergraduate Learning Outcomes. Poster presented at the Association of Public Policy Analysis & Management (APPAM) Annual Meeting. November 2018. Washington D.C.

Transferring from a Community College or Initiating in a Four-year College Directly: Estimating the Impact on Student Academic & Labor Market Outcomes. CC2PhD: UCLA Community College Studies Conference. June 2018. Los Angeles.

*EASE'ing Students into College: The Impact of Multidimensional Support f*or *Underprepared Students*. American Educational Research Association (AERA) Annual Meeting. April 2018. New York.

New Faculty Types in Higher Education: The Impact of Lecturers with Security of Employment on Undergraduate Learning. Association for Education Finance and Policy (AEFP) Annual Meeting. March 2018. Portland.

*EASE'ing Students into College: The Impact of Multidimensional Support f*or *Underprepared Students*. Society for Research on Educational Effectiveness (SREE) Annual Meeting. February 2018. Washington D.C.

Social-Psychological Interventions: Can They Improve STEM Persistence & Beliefs? Association of Public Policy Analysis & Management (APPAM) Annual Meeting. November 2017. Chicago.

What Works for College Student Success: A Meta-Analytic Study of Motivation Interventions. Poster presented at the Association of Public Policy Analysis & Management (APPAM) Annual Meeting. November 2017. Chicago.

Meta-Analysis of Random Assignment Studies of Interventions Targeting the Identified Competencies. Society for Research on Educational Effectiveness (SREE) Annual Meeting. March 2017. Washington D.C.

Looking Beyond Academic Performance: The Impact of Instructor Gender on Student Engagement and Attitude in STEM Fields. Association for Education Finance and Policy (AEFP) Annual Meeting. March 2017. Washington D.C.

Racial Segregation in California Higher Education: Examining Latino Student Population Trends Between and Within Sectors. Poster presented at the Association for Education Finance and Policy (AEFP) Annual Meeting. March 2017. Washington D.C.

College Majors: Their Role in the STEM Pipeline. Poster presented at the Population Association of American (PAA) Annual Meeting. April 2016. Washington D.C.

UNIVERSITY TEACHING EXPERIENCE

University of California, Irvine, School of Education

Teaching Assistant & Lab Instructor, Education 288A: Graduate Statistics I, 2015-2017
Teaching Assistant & Lab Instructor, Education 15: Statistics in Education Research, 2015-2017
Teaching Assistant, Education 10: Education Research Design, 2015

K-12 TEACHING & MENTORING EXPERIENCE

Arnold O. Beckman High School, Irvine, CA

Advanced Placement Microeconomics, September 2007 – June 2013

Early Academic Outreach Program (EAOP), UCLA

Academic Counselor, September 1998 – June 2001

RESEARCH EXPERIENCE

Solving the Equation: Recruiting, Hiring, and Retaining Math and Science Teachers Project, Research Assistant, 2016 – present
PI: Susanna Loeb

The National Academies of Sciences: Supporting Students' College Success: The Role of Assessment of Intrapersonal and Interpersonal Competencies, Consultant, 2016 - 2017
Chair: Joan Herman

SERVICE

University of California, Irvine, School of Education

First Beijing Normal University-UCI Doctoral Student in Higher Education Leadership and Management Exchange Program, Student Coordinator, 2017-2018
Faculty Search Committee, Student Representative, 2017-2018
Ph.D. Admissions Committee, Student Representative, 2015-2016
Diverse Educational Community and Doctoral Experience (DECADE), Peer Mentor, 2015-2018

Reviewer

American Educational Research Association Open
Journal of Research on Educational Effectiveness

SKILLS

Languages: English (native)
Statistical Programming: STATA (Advanced), Comprehensive Meta-Analysis (Intermediate), R (Beginner)

PROFESSIONAL MEMBERSHIPS

Association of Public Policy Analysis & Management (APPAM)
American Education Research Association (AERA)
Association of Education Finance and Policy (AEFP)

ABSTRACT OF THE DISSERTATION

Three Essays about Higher Education Programs and Policies
to Support Student Success

By

Sabrina Mahendra Solanki

Doctor of Philosophy in Educational Policy & Social Context

University of California, Irvine, 2019

Assistant Professor Di Xu, Chair

The goal of this three-study dissertation is to examine ways that institutions can support student success and college persistence with the goal of ensuring improved outcomes. Study 1 is a meta-analysis of theoretically-driven social-psychological interventions that aim to both improve the college experience for students and improve traditional markers of academic success such as course grades and college persistence. Overall, social-psychological interventions yield effects (d = 0.15) that are comparable to effects found in other interventions in the college context. Further, they are particularly effective for specific groups of students, such as students underprepared for college-level coursework and under-represented ethnic and racial minority groups. Study 2 is an evaluation of a two-year learning communities program for incoming biological sciences students at a large research university in California. The results suggest that the learning community program does indeed improve the student experience in college, as measured by *sense of belonging* and *academic integration*. Additionally, the learning community program impacts academic outcomes, affecting introductory first-year biology courses and

cumulative GPA, for example. Lastly, Study 3 focuses on college instructors, evaluating the relative impact of different types of faculty on student learning and engagement outcomes. I find limited evidence supporting differences by faculty type overall with one notable exception: students who took their introductory course with tenure-track teaching faculty earn slightly better subsequent grades and accumulate more subsequent credits than otherwise similar students who took their introductory course with lecturers or research faculty.

INTRODUCTION

In today's competitive labor market, a college degree is more valuable than ever before. Of the 11.6 million jobs created after the Great Recession, almost 66% have been secured by those with at least a bachelor's degree. Individuals with an undergraduate degree not only have a significantly greater chance of securing a job, but also earn roughly 35% more than those with only a high school diploma. Further, degree attainment has been shown to be a strong predictor of social mobility, overall health and well-being, and civic engagement (Baum & Payea, 2004, 2005; Baum & Ma, 2007; Shapiro et al., 2017).

However, college persistence remains a problem. While nearly 18 million undergraduate students were registered at institutions of higher learning in 2014 (Ginder, Kelly-Reid, & Mann, 2015), representing a 35% increase in enrollment since 2000 (Snyder & Dillow, 2015), the number of students receiving their bachelor's degree within a six-year period has increased only modestly within the same time period and remains quite low in 2019. It is troubling—yet perhaps not surprising—that as of today, only 60% of students who attend a four-year institution earn a degree. Further, even fewer first-generation students and under-represented ethnic and racial minority groups—44% of Black students and 55% of Hispanic students, to be exact—earn degrees, which is particularly concerning since fewer minority students enroll in college to begin with (Baum et al., 2015).

College persistence models have identified many factors contributing to low college persistence rates. In summary, these models contend that student academic progress in college—particularly during students' first year—is strongly correlated with student retention decisions and degree attainment. For example, Adelman (1998) found first-year college grades were positive predictors of degree completion. Similarly, Pascarella and

Terenzini (2005) found college grades to be one of the most consistent predictors of student persistence and degree completion. More recently, Gershenfeld, Ward Hood, and Zhan (2016) found academic performance—specifically in terms of low first-semester GPA—was a statistically significant factor in explaining why underrepresented students did not graduate within the 6-year time frame.

In addition to academic progress standards, college persistence models stress the need for institutions to focus on subjective student experiences when designing support programs. Rather than focusing solely on how to support students academically, research indicates that programs need to consider the psychological processes that serve as barriers to student learning. For example, some students find the academic environment alienating, feel out of place, and/or lack the confidence needed to succeed. To further an understanding of this idea, Robbins et al. (2004) examined the relationship between a number of psychosocial and study skill factors and college outcomes, such as cumulative GPA and college persistence. They found that the best predictors for GPA were academic self-efficacy and achievement motivation. Further, psychosocial and study skill factors were still predictive of both college outcomes even after socioeconomic status, standardized achievement, and high school GPA were controlled for.

Societal evidence and available literature concerning college persistence have made one idea very clear: persistence is informed by a variety of factors and thus extremely nuanced. As such, institutions must tackle the complexity of college persistence creatively and experimentally, the fact of which inspires the studies described below. The goal of this three-study dissertation is to examine strategies institutions can implement to support student success and college persistence with the goal of ensuring improved outcomes.

The three studies are largely informed by the college persistence models discussed above. They are motivated by prior empirical work documenting strategies that seem particularly effective at increasing course success and college completion rates, such as providing holistic student support and using evidence-based teaching methods in the classroom (Schneider & Clark, 2018).

I ask three distinct research questions:

1) What is the relationship between low-cost, light-touch social-psychological interventions and academic achievement for students in college?
2) Do learning communities improve the college experience as well as improve academic outcomes if implemented in STEM programs?
3) Do students learn similarly from different types of faculty?

In the remainder of this chapter, I briefly summarize three empirical studies in which I address the research questions identified above and consider the significance of this dissertation as a whole. In the subsequent three chapters, I fully describe each of these studies. I conclude this dissertation by discussing the implications of my findings and directions for future research in Chapter 4.

Study 1 is a meta-analysis of theoretically-driven social-psychological interventions that aim to improve both the college experience for students and traditional markers of academic success, such as course grades and college persistence. These interventions directly leverage student mindsets and motives to improve student learning. A number of narrative reviews suggest that these types of interventions are effective; however, to date, there is no meta-analytic review that focuses on these interventions in the college context. The interventions are conducted in college classrooms and laboratories at a rapid pace, which not only substantiates a focus on the college context, but also makes a meta-analytic review feasible.

The meta-analysis described in study 1 uses a strict inclusion criteria, focusing only on interventions that use randomized control design. Additionally, it follows the framework outlined in Harackiewicz & Priniski (2018) and focuses on three types of social-psychological interventions relevant to higher education. These include interventions that focus on students' perceived value of academic tasks ("task value interventions"), their framing of academic challenges ("framing interventions"), and their personal values ("personal value interventions"). The meta-analysis identifies 42 intervention studies.

Overall, I find that social-psychological interventions yield effects (d = 0.15) comparable to effects found in other college interventions. Further, I find that they are particularly effective for specific groups of students, such as students underprepared for college-level coursework and under-represented ethnic and racial minority groups. I also find, however, that these interventions have mainly been conducted at research-intensive universities and that the effect size of these institutions is much larger than that of broad-access schools and community colleges. I also report larger effects sizes for studies conducted in laboratories than for natural settings, such as a classroom. These two findings in particular call into question whether social-psychological interventions can be effective for the majority of students attending college today. I detail implications for higher education administrators, practitioners, and researchers in the discussion section, also considering the costs of these programs.

Study 2 is an evaluation of a two-year learning communities program for incoming biological sciences students at a large R1 university in California. The learning communities program, called EASE, is a way to provide holistic student support. It is informed by Tinto's (1975) model of student persistence and therefore includes forms of support that can help students become integrated in college both socially and academically. Prior research about learning communities has shown that learning community programs are associated with student academic achievement and retention; further, they improve the overall student experience particularly in the critical first year of college.

While these studies are informative, only a small quantity use research designs that address potential selection bias issues, and only a few were conducted in a STEM context. Study 2 addresses potential selection uses by using a regression discontinuity design based on the fact that students are assigned to the program if their SAT Math score falls below a certain threshold; it focuses on students in the biological sciences major at a large public university in California. I include a variety of student performance measures as outcome measures—including grades in gateway biology courses, first-year cumulative GPA, and retention within the biological sciences major by the end of the first year—as well as non-cognitive measures of social integration, such as sense of belonging within the biological sciences discipline at this school. Further, given that learning community components—such as the fostering of student-faculty and peer-peer interactions—are particularly beneficial for students typically underrepresented in STEM fields, study 2 also includes an analysis to explore treatment effect heterogeneity.

Results suggest that the EASE program does indeed improve the student experience in college. Students in the EASE program reported higher values for the measures *sense of belonging* and *academic integration*. Additionally, I find that the EASE program impacted academic outcomes, affecting introductory first-year biology courses and cumulative GPA, for example. It did not, however, have an impact on retention, as measured at the end of students' first year in college. Lastly, I did not find differential impacts for subgroup populations of students. Using EASE as a case study, the discussion section details implications for higher education administrators and practitioners in terms of promoting student success in STEM.

Study 3 focuses on the college classroom, evaluating the relative impact of different types of faculty on student learning and engagement outcomes. Study 3 is motivated by a line of inquiry suggesting that faculty play a central role in undergraduate education. It aims to shed light on the relative impact of different types of faculty on student outcomes and is therefore related to a small but growing body of literature about instructor effectiveness at the postsecondary level. This study is the first to include a new category of faculty never examined before—teaching-oriented faculty with tenure appointment. These are faculty whose primary responsibility is teaching rather than research. At a number of institutions, including the institution in the present study, teaching-focused faculty also engage in discipline-based education research and often serve as educational leaders within their departments. Their responsibilities and roles on campus thus make them well-positioned to improve undergraduate education. However, little is currently known about teaching-focused faculty, and more importantly, no study has linked teaching-focused faculty to student outcomes. Given a longstanding debate about whether institutions should grant tenure to teaching-faculty, knowledge about their impact on student academic outcomes relative to other types of faculty provides an empirical foundation for better-informed staffing decisions made by policymakers and college administrators.

Using data from six cohorts of students at a single UC campus, I examine the impact of initial course-taking with three distinct types of instructors—tenure-track research faculty, tenure-track teaching faculty, and lecturers—on students' current and subsequent academic outcomes. Using a three-way fixed effects model, I find overall limited evidence supporting differences in faculty type with one notable exception: students who took their introductory courses with tenure-track teaching faculty earned slightly better subsequent grades and accumulated more subsequent credits than otherwise similar students who took their introductory courses with lecturers or research faculty. These results provide several implications for institutional faculty hiring policies and inform promising avenues for future research, which are detailed in the discussion section.

The studies outlined above concern pressing issues in higher education today and provide insight about effective support mechanisms and institutional structures that help students cross the finish line. It is my hope that they inform higher education administrators and practitioners and can be used to promote opportunity and improve learning outcomes for all students.

CHAPTER 1

Charting College Success:

A Meta-Analytic Study of Social-Psychological Interventions

1. Introduction

A college degree is crucial in today's competitive labor market, yet nearly half of all college students never graduate (Shapiro et al., 2017). Because of statistics such as this, institutions of higher education, in addition to policymakers at different governmental levels, have made college completion a national priority. Researchers, policymakers, and practitioners seek to better understand which factors are critical to student success at different points in students' college-to-degree trajectory. What is known today is that the college completion puzzle is complex, including factors such as college readiness, institutional support, non-cognitive characteristics, financial considerations, and job opportunities.

Recently, attention has been paid to the powerful role social-psychological factors—such as student mindset and motives—can play in helping encourage college student learning, success, and degree completion (Destin, 2018; Harackiewicz & Priniski, 2018). Indeed, a number of correlational studies have concluded that social-psychological factors are critical to success in the college setting (Chemers, Hu, & Garcia, 2001), linking these factors to student engagement and positive behavior. In addition, studies have found that these factors predict both academic achievement and college persistence (Robbins et al., 2004; Robbins, Allen, Casillas, Peterson, & Le, 2006). Further, a number of interventions based on social-psychological and motivation theories have been developed over the past two decades in an effort to improve educational outcomes in higher education. Though all *narrative* reviews about these interventions are positive (see Harackiewicz & Priniski, 2018; Karabenick & Urdan, 2014; NASEM report, 2017; Yeager & Walton, 2011), the evidence from systematic reviews, such as meta-analysis, is limited and includes mixed results. Additionally, the majority of systematic reviews estimate the effect of interventions across the K-20 school spectrum rather than focusing on higher education specifically. Thus, it remains relatively unclear whether social-psychological factors are useful points of intervention for researchers and practitioners to use to actively promote student success *in college*.

The present study fills this gap in research literature by examining whether approaches to college completion that incorporate social-psychological factors—namely, interventions that encourage a growth mindset, link classroom work to real-world aspirations, or use tools to activate students' motivation and sense of belonging—typically improve student success in higher education. Different from the narrative reviews that dominate education literature, this study uses a meta-analytic approach. The rapid growth of college motivation interventions over the past few years renders a meta-analysis in the college setting feasible, and substantiates a focused review of this setting. In our analysis, we include interventions that use randomized control trials—the gold standard in research design. The present study will therefore quantify the average causal effect of these interventions, which could be informative for institutional administrators who must decide which student success policies to support.

The present study also contributes new knowledge to the body of meta-analytic studies in the field. First and foremost, it includes more studies than the most recent review conducted on this topic (see Lazowski & Hulleman, 2016). The additional data points help us arrive at a more precise estimate of intervention effects and can also provide evidence about the degree to which intervention effects can be replicated. The impacts of social-psychological interventions are very much context-dependent (Yeager & Walton, 2011). In fact, there have been calls for more research about moderator variables and context that have yet to be examined, which the present study includes. Specifically, researchers note that what might work in a research-intensive university with high-performing, highly-motivated students may not be as effective at a more broad-access, less selective school, for example. Also, challenges students face in which these interventions aim to help overcome—such as stereotype threat—could be more pronounced at certain institutions where the proportion of under-represented minority students (URM) and low-income students is small.[1] Researchers therefore consider it important to understand where effects are generated, and this is particularly important for institutional administrators and practitioners so that they are not compelled to implement policies based on interventions not appropriate for their student population. As such, in the present study, we include a variable to examine this matter further. This variable identifies whether the study was conducted at a selective research university, less selective university, or broad access university. The study also includes a variable that identifies whether the institution is public or private.

Lastly, we also include moderator variables that reflect both short and long-term achievement measures and student demographic characteristics. This study therefore will

[1] Stereotype threat refers to the anxiety one can experience as a member of a stigmatized social group, particularly when asked to engage in a task related to the negative stereotype associated with the social group (Steele, 1997). Exposure to stereotype threat is hypothesized to be a reason females and minority students disengage in certain fields such as STEM.

shed light on whether these interventions are effective for subgroup populations of students, such as URM students, in addition to whether these interventions have enduring impacts. Overall, the present study has the potential to facilitate a greater understanding of these interventions, thereby positively informing the decisions of institutions, educators, and policymakers.

2. Social-Psychological Interventions in Higher Education

The college context is a unique setting for motivational interventions. College is a time of transition that has the potential to amplify feelings of self-doubt, anxiety, and isolation (CCMH report, 2018; Eagan et al., 2016). These feelings, coupled with increased personal autonomy, make college students well-positioned to reap the benefits that social-psychological interventions can offer. Social-psychological interventions are based on social-psychological and motivation theories, targeting specific educational problems—such as the challenges students face during their first year in college—and the processes underlying them. For example, in one study, students were provided with opportunities to make concrete connections between what they were learning in a course and things that they cared about. This particular intervention increased final exam scores in the course, and the effects were most pronounced for students who performed poorly on their initial exams (Hulleman et al., 2017). In another study, students were given an interactive module to watch during college orientation that explicitly conveyed the idea that intellectual ability can change and develop (i.e., a growth mindset). Students exposed to this module significantly improved their academic motivation and performance (Yeager et al., 2016).

The present study focuses on three types of social-psychological interventions relevant to higher education, following the framework put forth by Harackiewicz & Priniski

(2018). These interventions focus on students' perceived value of academic tasks ("task value interventions"), their framing of academic challenges ("framing interventions"), and their personal values ("personal value interventions") (Harackiewicz & Priniski, 2018). In the section that follows, we provide examples of interventions that fall into each category.

2.1 Theoretical Categories of Social-Psychological Interventions in College

Task value interventions communicate the value and/or importance of course content either by providing examples of the relevance or usefulness of academic tasks for personal goals, or by encouraging students to ascertain task value for themselves through writing exercises (Harackiewicz & Priniski, 2018).

Utility value interventions are a common type of task value intervention in higher education. Utility value interventions typically involve asking students to write about why a particular aspect of their course was relevant or useful. For example, researchers in Canning et al. (2017) asked undergraduate biology students in the experimental group to write essays addressing the personal relevance of a topic from their course. They were encouraged to include concrete information covered in the unit and to explain why this information was relevant and useful, using personal examples. The experimental group earned significantly higher grades (on a 0 to 4 point scale) in a biology course compared with their peers in the control group (M = 2.96, SD = .71 vs. M = 2.77, SD = .75), and 79.8% of students in the experimental group proceeded to the next semester of biology, as compared to only 69.5% in the control condition.

In another example, Harakciewicz et al. (2015) implemented a brief writing intervention in a college biology course. Students in the treatment group were asked on three separate occasions to answer a question using course material and to discuss the relevance of the concept or issue to their own life or the lives of others. Those receiving treatment saw a significant—but relatively small—overall improvement in biology course grades (d = .06). The effect was particularly pronounced for first-generation URM students whose grades rose by more than half a letter grade, on average, compared with the grades of the control group (slightly over 2.5 versus slightly under 2.1 on a 0 to 4 point scale).

Framing interventions focus on the challenges students face during academic transitions and help students cope with adversity by encouraging them to frame challenges as common and surmountable (Harackiewicz & Priniski, 2018).

Social belonging interventions are a type of framing intervention. Interventions targeting social belonging—that is, a feeling of connectedness and relatedness to one's institution and peers—typically involve making students aware that uncertainty about their place in college is both common and temporary. Walton, Logel, Peach, Spencer, and Zanna (2015) utilized a social belongingness intervention targeting freshman engineering students. Participants in the intervention group read a brief report and listened to audio recordings of upperclassmen who confirmed that both male and female engineering majors worried about belonging and representation in engineering, but experienced the fading of these feelings with time. The intervention resulted in impressive treatment effects, as the GPA of women in male-dominated majors rose, on average, more than a full letter grade higher than the GPA of otherwise similar women in the control group. In addition to the effect on GPA, this intervention improved women's perceived experience of their engineering major—as measured by a sense of belonging, self-efficacy, and enjoyment—relative to the control condition (a moderate-to-large effect size equal to a standardized mean difference of .67).

Growth mindset interventions also fit this category. A growth mindset is a student's belief that intelligence is not a fixed entity and is instead a malleable quality that can be improved. Mindset theory suggests that students with stronger growth mindsets have more adaptive psychological traits and behaviors (e.g., a positive response to failure), which lead to greater academic achievement (Dweck, 2000; Rattan, Savani, Chugh, & Dweck, 2015). Studies suggest that students with a fixed mindset can be taught that intelligence is in fact malleable and therefore can be improved with education. For example, Aronson, Fried, and Good (2002) used a pen-pal paradigm intervention in which students were encouraged to view intelligence as malleable. Participants were presented with fictitious letters from at-risk youth; those in the experimental condition were asked to write responses that would encourage their pen pals to work hard even though they were experiencing difficulties. They were further encouraged to talk about how research showed that the brain is like a muscle that will grow stronger through effort. Not only did the intervention measurably improve growth mindset beliefs, but it also improved educational enjoyment and GPA.

Attribution retraining interventions are similar, conceptually, to mindset interventions, as their goal is to link success to effort rather than to inherent ability. Theoretically, attributional retraining should engender the belief in students that they can succeed if they try harder or perhaps study more effectively (Menec, 1994). In a typical attribution retraining intervention, researchers in Perry, Stupnisky, Hall, Chipperfield, and Weiner (2010) presented college students with the message that poor performance can be improved through effort, accompanying this message with an activity as reinforcement. This study resulted in statistically and practically significant treatment effects, with attribution-retraining participants earning first-year GPAs that were approximately .26 grade points higher than those of control participants.

Personal values interventions focus on students' core values. As described in Harackiewicz & Priniski (2018), these interventions require students to articulate their personal values rather than merely contemplate the value of course material. This approach aims to highlight the indirect value of a course. For example, students may be asked to select from a list of values (e.g., independence, creativity, relationships with family) and then be asked to elaborate on why those values are important to them.

As one example of this intervention type, Harakciewicz et al. (2014)'s values affirmation intervention targeted students taking an introductory biology class. The intervention was delivered through a brief writing exercise, used twice during the semester, in which students wrote about the values most important to them. Control group students wrote about why values least important to them might be important to someone else. The effect of this type of intervention on GPA was significant for first-generation students but not for continuing-generation students. Further, the intervention promoted continued enrollment in the biology course sequence; first-generation students in the values affirmation group were more likely to enroll in the next biology course than continuing-generation students, representing a difference in enrollment of roughly 10%. They were also more likely to enroll in the next course in a relevant sequence than were their counterparts in the control group, representing a difference in enrollment of roughly 20%.

Another example of this type of intervention study is Brady et al. (2016). In this study, first and second-year college students completed a brief writing exercise in which those in the affirmation condition wrote about their most important values and control group students wrote about a value relatively unimportant to them, also exploring why this value might be important to someone else. Brady et al. (2016)'s intervention effect size was positive and moderate—but only for Latino students. For Latino participants, though, the benefits persisted: GPAs collected 2 years after the end of the intervention were significantly higher for the affirmation condition than those of their counterparts in the control group. This affirmation intervention also resulted in a large (effect size equal to 0.94) positive impact for Latino students regarding "adaptive adequacy"—a measure of self-integrity, self-esteem, and hope.

2.2 Important Features of Social-Psychological Interventions in College

The studies outlined above indicate that social-psychological interventions are theoretically different from one another and have effects that are often more pronounced for certain subgroup populations of students. They also demonstrate that these interventions show promise in regard to helping students perform better in their courses. It is also important to note that these interventions differ in regard to their features, such as design elements. These differences, discussed below, have the potential to moderate the relationship between interventions and academic achievement.

Design elements. Interventions targeting motivation differ with regard to their design elements. Two main differences are (1) whether an intervention includes an interactive component and (2) intervention setting.

Passive versus interactive intervention. In regard to motivation interventions, an interactive element means participants read materials about the targeted motivation factor and then write a relevant reflective essay. A writing component has the potential to enhance treatment effects because it can help students internalize positive messages about success through a 'saying is believing' effect (Higgins & Rholes, 1978; Aronson, 1999). Engaging with messages in this way has greater potential to foster an attitude change than passively receiving a message alone.

Intervention setting. Motivation interventions can take place in a variety of settings, such as laboratories or in a classroom. While a laboratory setting can promote internal validity by controlling for extraneous factors, laboratory interventions run the risk of lacking ecological validity. In fact, Lazowski and Hulleman (2016) found that average intervention effect sizes dropped from .63 to .46 when only studies conducted in the educational context (i.e., excluding laboratory-based studies) were considered, suggesting that what is effective in the laboratory may be less effective in contextualized settings. In contrast, classroom-based interventions are more difficult to control but have the advantage of being situated in the learning context—findings from these studies therefore have greater potential to support the external validity of social-psychological interventions, as compared to studies conducted in a lab.

2.3 Review of Previous Narrative Reviews and Meta-Analyses

Narrative reviews. A number of reviews discuss the value of motivation interventions. Only a small number of them, however, focus on the college setting. As an example, Karabenick & Urdan (2014) discussed the importance of empirical evidence in motivation interventions, examining attribution retraining, expectancy-value theory, and identity and interest theories. The numerous studies and projects outlined in this research synthesis further emphasize the effectiveness of motivation interventions in promoting positive student outcomes—such as improved academic performance and persistence—at all grade levels. One particularly salient takeaway from this review is that motivation interventions are most effective when they target a specific group of students at a specific point in their educational careers. This reasoning substantiates our focus on college-level interventions, as college is a unique educational period in which many students experience autonomy for the first time, face new academic and social opportunities, and must manage a new set of responsibilities. These changes have the potential to alter students' motivation for learning and therefore require institutions to incorporate engagement techniques that are different from those used in the K–12 context.

A recent narrative review provided further support for motivation interventions (NASEM report, 2017), and focuses only on higher education. The report highlighted the role of intrapersonal competencies in students' college success, including qualities such as a strong sense of belongingness, growth mindset, and academic self-efficacy. As part of this effort, the report reviewed interventions that have helped students develop these competencies, reporting that interventions based on these qualities seem particularly effective for students most at risk for college attrition, such as URM students. NASEM concluded by encouraging replication of the interventions.

In a cross-grade narrative review, Yeager and Walton (2011) emphasized the theoretical underpinnings and efficacy of motivation interventions. It reviewed laboratory-based motivation intervention studies that communicate a social-psychological message, utilize random assignment, and estimate effects of interventions on students' grades in a course or institution over time. The authors concluded that such interventions are effective because they affect how students learn and thereby result in recursive self-reinforcing messages.

A more recent review of theory-based motivation interventions comes from Harackiewicz and Priniski (2018). In this review, the authors discussed interventions that aim to change students' perceived value of academic tasks, re-frame students' interpretation of academic challenges, and help students tap into sources of self-worth. The scope of the review resulted in the analysis of 20 studies, though the authors only discussed a subset in detail.

A contribution of this review is that it provides a theoretical discussion about the relationship between intervention type and student learning outcomes. For example, utility value interventions engage students with the content of a course and the content of a particular field. In other words, they have the potential to stimulate interest in a field and therefore could very well be related to more distal outcomes, such as students' educational and career choices. Echoing previous work, Harackiewicz and Priniski (2018) concluded that theoretically-driven motivation interventions are effective because they target specific educational problems and the processes underlying them. Further, they stressed that these interventions are context-dependent and encouraged replication to better understand the mechanisms with which and conditions under which these interventions are most effective. They also noted the importance of replication given that some interventions have inconsistent findings across studies, particularly in regard to personal values interventions.

Meta-analytic reviews. The narrative reviews discussed above are informative. Each review treats the empirical context, sample, and treatment conditions delicately, and the reader is provided with a detailed account of interventions and their potential benefits in regard to student learning outcomes. More often than not, however, narrative reviews include only a small subset of studies in a field, and they are therefore criticized for having an element of selection bias (Cooper & Rosenthal, 1980; Lipsey & Wilson, 2001). As compared to narrative reviews, systematic reviews encompass all studies that are publicly available. They also have the benefit of providing a meta-analytic effect size measuring the strength of the intervention effect. And, as mentioned earlier, this type of information is particular important for institutional administrators and policymakers who have limited resources and must decide between a number of approaches for supporting student success in college.

Sisk, Burgoyne, Sun, Butler, and Macnamara (2018) provide two meta-analyses focusing solely on the relationship between growth mindset and academic achievement across the K-20 school spectrum. Meta-analysis 2, most relevant to the present study, estimated the overall impact of growth mindset *interventions* on academic achievement, as measured by standardized test scores, course grades, and GPA. They concluded that overall effects were weak; on average, the academic achievement of students receiving a growth-mindset intervention and students in control groups was d = .08. Even though the meta-analysis included students in various developmental stages—that is, children (i.e., primary school students), adolescents (i.e., middle school, junior high school, and high school students) and adults (i.e., postsecondary students)—average effects were found to be similar across groups, hovering around .08. However, subsequent moderator analyses showed that socio-economic status (SES) was a significant moderator. For students from low-SES households, academic achievement was significantly higher for those who received the growth-mindset intervention, relative to control group students (d = .34, 95% CI = [.07, .62], p = .013). This particular finding echoes the narrative reviews discussed above. Lastly, Sisk et al. (2018) also found that mode of intervention was a significant moderator. Specifically, growth-mindset interventions were not effective when administered via computer programs, as compared to interventions administered via reading materials.

The Lazowski and Hulleman (2016) meta-analysis is the only quantitative review to date that includes a comprehensive dataset of theoretically-driven motivation interventions. Lazowski and Hulleman (2016) encompassed 92 intervention studies and 16 related theoretical frameworks of student motivation spanning all educational stages.

Because Lazowski and Hulleman (2016)'s inclusion criteria is less conservative than that included in our study, there is a wide breadth of utilized studies, and it is therefore difficult to determine the implications of this meta-analysis for higher education. Indeed, it reported an effect size of d = .47 (95% CI [.38 to .57]) for studies conducted in the postsecondary context. This effect size, however, included both randomized control trial studies and quasi-experimental studies; we cannot interpret the effect size estimate as a causal effect. The reported effect size also reflected both academic (e.g., course grade-related) and non-academic (e.g., self-efficacy-related) outcome measures. To what degree these interventions are effective for bottom-line college outcomes, such as course grades and GPA, is difficult to determine from this meta-analysis.

Further, the difference in effect size estimates between the Lazowski and Hulleman (2016) and the Sisk et al. (2018) meta-analyses suggest an updated analysis could be informative. Lazowski and Hulleman (2016) reported an effect size equal to .56, whereas Sisk et al. (2018) reported an effect size equal to .08[2]. Even accounting for the differences in inclusion criteria, the large disparity in effect size estimate that exists for growth mindset

[2] The effect size for Lazowski & Hulleman (2016) (d = .56) refers to growth mindset studies. See Table 2 in Lazowski & Hulleman (2016).

interventions between these two studies suggests that the additional studies included in Sisk et al. (2018) are mainly responsible for their different conclusion. This emphasizes the need to examine whether effect size estimates for other intervention types, such as utility value and social belonging, experience the same type of decline in effect size estimate with additional studies. We are able to address this in our analysis.

Though Lazowski and Hulleman (2016)'s meta-analysis is relatively new, the growth in interventions of this type since their last search in May 2015 has substantially increased the number of studies we analyze in the present study. The additional studies allow us to analyze the data further and include statistical contrasts that extend beyond the study's research goals. Specifically, we evaluate whether effect size estimates differ by outcome measure, study context, and subgroup populations of students.

2.4 Present Study

Existing literature about social-psychological interventions suggests that they have the potential to improve student learning outcomes. The present study enhances this body of work, contributing to current discussions about social-psychological interventions and their overall efficacy by providing a targeted perspective and an up-to-date meta-analysis of their impacts on college students' academic achievement. It includes 42 college-level randomized interventions (with 113 unique effect size estimates). We focus on randomized control designs since they are inarguably the gold standard in research design and best suited to address questions of cause and effect (Shadish, Cook, & Campbell, 2002).

Further, the present study investigates factors—such as type of intervention, study context, outcome measure, and intervention design features—that may moderate the relationship between social-psychological interventions and academic achievement. We focus on the following two research questions: (1) can social-psychological interventions improve college students' academic achievement? and (2) what characteristics of interventions and participants best predict differences in effect sizes across studies?

3. Method

3.1 Literature Search

A comprehensive literature search strategy was implemented using *Google Scholar* and *PsychInfo* to locate peer-reviewed journal articles, conference papers, doctoral dissertations, and other published and unpublished materials ('grey literature'). We retrieved studies published through the end of June 2019. We used the following search terms: *motivation interventions, motivation interventions college, growth mindset interventions, sense of belonging interventions, utility value interventions, attribution retraining, academic self-efficacy interventions, goal-setting interventions, positive future-self, positive future-self interventions,* and *intrinsic motivation interventions.* We also reviewed the reference lists of pertinent chapters in Karabenick and Urdan (2014), prior educational reviews in the same field (e.g., Lazowski & Hulleman, 2016; Yeager & Walton, 2011), and relevant empirical studies found in the database search. We also contacted authors known in the motivation field to request access to their unpublished work. Our literature search yielded 553 results in total.

3.2 Inclusion Criteria

Results were only eligible for inclusion if they reported empirical results from an intervention (i.e., they were a study). Intervention studies meeting the following eight criteria were included in this meta-analysis:

(1) The intervention targeted students attending a college or university,

(2) The intervention that the study presents aimed at directly and/or indirectly manipulating students' growth mindset, sense of belonging, utility value, and values affirmation,

(3) The study included a clearly defined treatment and control/comparison group,

(4) The study used random assignment,

(5) The study included an academic outcome variable related to college persistence,

(6) The study reported sufficient information so that effect sizes could be calculated at posttest,

(7) The study was published in English, and

(8) The study was published after 1980, up to 2019.

These inclusion criteria reduced the number of qualifying studies to 42 (see Figure 1.1), from which we computed 113 effect sizes (see the Calculation of Effect Size section for details).

3.3 Coding Procedure

A code sheet and codebook were developed to record information about effect sizes and potential moderator variables. Coded variables included (1) type of intervention (i.e., utility value, social belonging, growth mindset, values affirmation), (2) outcome measure, (3) institutional context, (4) design elements, and (5) publication status. We also included an indicator for whether the effect size reflected estimates for a particular subgroup population of students. The first and second authors of the present study independently coded each study, and the coding agreement across all dimensions was 97%. All disagreements were resolved via discussion.

Intervention type. Each study was coded following the classification system outlined in Harackiewicz and Priniski (2018), where targeted interventions in higher education are divided into three types: task-value, framing, and personal values. Using these distinctions and the examples provided for each (see discussion above Theoretical Categories of Social-Psychological Interventions in College), we further categorized studies in our analysis into the following four categories: (1) utility value, (2) social belonging, (3) growth mindset/attribution retraining, and (4) values affirmation.

Outcome measure. Each individual effect size was categorized as measuring one of the following four student outcomes: (1) competency test (e.g., assessments of critical thinking), (2) course exam, (3) course grade, (4) GPA, and (5) persistence measure.[3]

Study context. We coded each study for whether it took place at a public or private institution. We also coded each study in regard to whether the intervention was conducted at a selective institution. Selectivity includes two categories based on the institution's acceptance rate and the institution's 4-year graduation rate.[4]

Design elements. Interventions were coded for whether or not they included a writing component. Interventions were also coded for their setting; that is, whether they were administered in a lab, course, or online.

Student demographic characteristics. Individual effect sizes were coded for whether they were gathered only for Black students, Hispanic/Latino students, first-generation college students, or women in a male-dominated major.

---

[3] The following outcomes were categorized as a measure of persistence: first-year retention, retention at year 2, number of courses in a similar field, total credits accumulated during the study window, enrollment in subsequent course.

[4] See Table 1.2 for acceptance rates and graduation rates and corresponding selectivity category. We also use a selectivity variable that comprised of three categories. The conclusions made about selectivity and average ESE remains the same.

Publication status. We coded the variables related to publication status as follows: (1) the study's year of publication and (2) whether it was published in a peer-reviewed journal (1= published; 0 = unpublished).[5]

3.4 Calculation of Effect Size

Hedges' *g*, which adjusts Cohen's *d* to correct for upward bias for small samples (Hedges, 1981), is the standardized effect size (ES) used in our meta-analysis, to indicate the difference in outcome variable (i.e., academic achievement) in standard deviation units between the treatment and control groups.

ES values were calculated using the Comprehensive Meta-Analysis (CMA) software program (Borenstein, Hedges, Higgins, & Rothstein, 2005). These values are based on a variety of statistical information provided by each study: means, standard deviations, *p*-values, and *t*-values, for example. Positive ES values indicate a more favorable result for students receiving the intervention than for students in the control group. When a study did not provide enough information to calculate the ES, we contacted the authors for the appropriate statistics. In all, we contacted five authors, and two responded to our request.

3.5 Publication Bias

Publication bias occurs when some results are systematically less likely to be published than others (e.g., studies that find small or null effects; Rosenthal, 1979). We investigated publication bias in the following ways. First, we compared effect sizes from published studies to those from unpublished studies. The results are shown in Table 1.1 (Column 3). The mean ES estimate is .17 and .11 for published and unpublished studies,

[5] We did not find non-peer-reviewed published studies that satisfied our inclusion criteria.

respectively. The difference in these estimates, however, is not statistically significant (p = .48).

Second, we also investigated publication bias visually, by examining a funnel plot for any asymmetries in ES distribution, as well as statistically, by conducting an Egger's regression test for publication bias (Egger, Smith, Schneider, & Minder, 1997). Funnel plots are scatterplots of ES estimates plotted against a measure of study size (e.g., standard error). Asymmetric funnel plots indicate potential publication bias (Harbord & Harris, 2009). The funnel plot in Figure 1.2 shows symmetry at the top of the funnel, but less towards the middle and bottom, ESs appear to be positively skewed. Visually, there appears to be a degree of publication bias. The Egger's regression test (1997) is a regression-based method that provides a p-value associated with publication bias. In our case, the p-value is significant. Taken together, the distribution of ESs supports an inference of publication bias; studies with smaller sample sizes are associated with more positive ESs than one would expect by chance.

3.6 Analytic Strategy

To estimate mean ES estimates, we use multilevel modeling procedures outlined in Lee, Warschauer, and Lee (2018), Pratt, Turanovic, and Cullen (2016), and Pratt, Turanovic, Fox, and Wright (2014). This strategy is appropriate given that our sample of effect size estimates is nested within a two-level hierarchical structure—that is, effect size estimates (level 1) are nested with studies (level 2). To execute this procedure, we used Stata 14 (StataCorp, 2015) and the command, *meglm*. In our specification, we included the sampling variance (i.e., squared standard errors) of the ES estimate in the random part of the model (i.e., level 1).

Specifically, we estimate the following equations:

$$Level\ 1: \qquad ESE_{ij} = \zeta_{ij} + e_{ij} \tag{1}$$

$$Level\ 2: \qquad \zeta_{ij} = \gamma_{00} + \upsilon_{0j} \tag{2}$$

At level 1, ESE_{ij} is the observed ES i in study j. ζ_{ij} is an estimate of the "true parameter" value of the ES i in study j, and e_{ij} is the sampling error of the ES i in study j. The e_{ij} term has a known variance of σ_{ij}^2. At level 2, υ_{0j} is the study-level random error. In combining the level 1 and level 2 equations, we obtain the unconditional model:

$$ESE_{ij} = \gamma_{00} + \upsilon_{0j} + e_{ij} \tag{3}$$

where the intercept γ_{00} is the mean ES estimate.

After computing an overall mean ES to assess the relationship between social-psychological interventions and academic achievement in college, we then conducted a series of moderator analyses to determine the degree to which intervention characteristics (such as type of intervention and design element) and characteristics of study participants impacted the overall mean. This entailed calculating the mean ES estimate for each coded variable described in "Coding Procedure" above.[6] Specifically, we included the moderating variable (for example, type of intervention) in equation (3).

4. Results

4.1 Overall Mean Effect Size Estimate (ESE)

Figure 1.3 presents a histogram of 113 unique ESEs used in this study; as shown in

[6] As detailed in Coding Procedure, we coded effect size estimates for whether they were gathered for only Black students, Hispanic/Latino students, first-generation college students, or women in a male-dominated major. These effect sizes were grouped together for the moderator analyses to determine whether effect size estimates differed for students at risk in college or major. For this specific analysis, we only include effect size estimates if provided for the at risk group and its' counterpart (for example, Black students and white students).

this figure, the ESEs range from -.88 to 1.45.[7] This figure suggests a positive relationship between motivation interventions and academic achievement, as the majority (70.8%) of these estimates show that students in treatment groups outperformed students in control groups. The red line indicates the overall mean ESE equal to .15.

As discussed earlier, our Egger's test of publication bias suggests that this form of bias may indeed be present in the literature about social-psychological interventions. As a result, we calculated an adjusted overall average ES to account for publication bias using the "trim and fill" approach (Duval & Tweedie, 2000a, 2000b). The trim and fill approach builds on the key idea behind the funnel plot: in the absence of publication bias driven by statistical significance standards, the funnel plot would be symmetric around the meta-analytic mean effect. If the plot contains more small studies on the right than on the left, this may be due to publication bias, where smaller samples were only published if they reported the larger effects required to achieve statistical significance. Conversely, the concern is that studies may be missing on the left, where non-significant studies were suppressed from the literature. The trim and fill method attempts to impute ESs into a funnel plot to achieve symmetry, adds these ESs into the meta-analysis, and then re-computes the meta-analytic mean (Borenstein, Higgins, & Rothstein, 2009). This mean based on trim-and-fill falls to .08 but is still statistically significant (CI ranges from .04 to .13).

4.2 Moderator Analyses

In addition to the overall mean ESE, Table 1.1 Column 3 shows the mean ESE for each moderator variable. Each estimate was calculated using bivariate multi-level models

[7] See Table 1.3 for a list of studies included in the analysis and the respective study-level ESE.

(i.e., models containing one independent variable). As shown, all ESEs are positive, and most are significantly different from zero.

In terms of intervention type, we find that utility value and growth mindset interventions are particularly effective. The mean ESE is .20 and .18 respectively. The mean ESE is lowest for studies categorized as 'values affirmation' (d = .05). As shown in Table 1.1, we also estimate the mean ESE for each outcome measure. The mean ESE is largest for those generated from competency test scores (d = .31) and GPA (d = .17), the latter of which includes the largest number of ESEs in the academic outcome measure category. A joint test of significance shows that the ESEs are marginally different from one another (Column 5, p = .08).

We also calculated mean ESEs for variables related to study context. We find that institutional context matters. Specifically, studies conducted at institutions coded as selective have a mean ESE equal to .18. Studies conducted at broad-access institutions have a mean ESE equal to 0.01. A joint test of significance shows that the ESEs are statistically different from one another (Column 5, p = .03).[8] Also in regard to study context, we calculated the mean ESE for studies that took place in a lab. The mean ESEs for studies that took place in a laboratory (d = .23) are much larger than those for studies that didn't take place in a lab (d = .11); this difference is statistically significant at the $p < .10$ level (column 5, p = 0.07).

With regards to the design elements variable, we find that counter to prior conclusions, studies including a writing component (d = .12) report average ESs similar to

[8] When institutional selectivity includes three categories—selective, moderately selective, broad access—the mean ESE is still larger among selective and moderately selective institutions as compared to broad access institutions. The joint test of significance p-value in this case is equal to .07.

those that do not include a writing component (d = .14). Lastly, we also calculated mean ESE for a subset of effects that reflected students at risk in college or within a major. The mean ESE for this particular group is .26 which differs significantly from the mean ESE calculated for a subset of effects for students not at risk (d = .04).

5. Discussion & Conclusion

In this study, the primary goal has been to provide evidence regarding the efficacy of social-psychological interventions in the college context. Our choice to focus on college-level interventions stems from the fact that college is a unique setting in which students become particularly autonomous. Unlike primary school and high school, during which parents and teachers often drive and incentivize students' motivation to succeed, college facilitates student responsibility and motivation. This notion, informed by the recent growth of social-psychological interventions in the college context, drove our inquiry.

Narrative reviews (Harackiewicz & Priniski, 2018; NASEM report, 2017) and the most recent systematic review (Lazowski & Hulleman, 2016) suggested that social-psychological interventions are an effective way to improve student academic outcomes in college. Consistent with this body of work, we found a positive relationship between this type of intervention and academic achievement, and as shown in Figure 1.3, a majority of ESs are positive and significantly larger than zero. It is important to note, however, that although our results are consistent in *direction* with extant knowledge, our ESE is one-third the magnitude of Lazowski and Hulleman (2016). This is not surprising when one considers the differences in inclusion criteria between Lazowski & Hulleman (2016) and the present study. As mentioned earlier, Lazowski & Hulleman include 16 different types of interventions. We include only experimental designs of four that are most relevant to the college context: utility value, sense of belonging, growth mindset, and values affirmation (Harackiewicz & Priniski, 2018).

Additionally, unlike Lazowski & Hulleman (2016), we include only academic-related outcome measures, excluding all non-academic outcomes—such as socio-emotional measures—in calculating our overall mean ESE. We speculate that the impact of ESs on non-cognitive outcomes is, in many cases, positive and significant, thereby increasing Lazowski and Hulleman (2016)'s overall average. Indeed, changes in non-cognitive measures are the likely mechanisms through which change in academic outcomes are achieved, and it is generally easier to move the intermediate outcome then the final one (Ruiz-Primo, Shavelson, Hamilton, & Klein, 2002). Lastly, given our more narrow focus on college-level *randomized* interventions, we include a small subset of Lazowski and Hulleman (2016)'s studies that fall within the four types of interventions mentioned above and include them in our review. Also, given the increasing popularity of these types of interventions, we include 31 additional studies unavailable to Lazowski & Hulleman (2016).

Different from the meta-analytic and narrative reviews we refer to in this study, the present study answers a more specific question: are social-psychological interventions effective in postsecondary education in terms of academic outcomes? We acknowledge that academic outcomes are not the only markers of success in college. For example, Brown et al. (2015) found that an intervention emphasizing the communal utility value of biomedical research increased student motivation to pursue a biomedical research career in the future. This is an important study with practical implications. We do not capture outcomes such as this in our analysis. We focus on bottom-line academic outcomes because they are highly relevant to institutional administrators and policymakers who are trying to understand the college to degree pipeline. We know from prior literature that how well students do academically is a strong predictor of college degree attainment (Gershenfeld, Ward Hood, & Zhan, 2016; Stewart, Lim, & Kim, 2015).

In our synthesis, we find that social-psychological interventions in college improve academic outcomes with a mean ESE equal to .15. The magnitude of the overall ESE we report is moderate (Kraft, 2018) and comparable to ESs often found in higher education research. For example, the Niu, Behar-Horenstein, and Garvan (2013) meta-analysis of instructional interventions aimed at improving college students' critical thinking skills reports an overall average ES of .20 for 40 intervention studies. In another recent example, Sneyers and De Witte (2017) examined mentoring interventions in higher education through a meta-analytic review of 25 studies, finding that student-faculty mentoring has a positive impact on both retention and graduation (d = .15 and .10, respectively). Comparing the average ES found in the present study to practical benchmarks in higher education may provide incentive for colleges and universities to implement social-psychological interventions and test their efficacy, especially since they are cost-effective. In fact, a typical social-psychological intervention costs only dollars per student (Yeager & Walton, 2011, Paunesku, 2013).

5.1 Key Findings from Moderator Analyses

We want to highlight two key findings from our moderator analyses that have clear implications for institutional administrators and future research. Social-psychological interventions seem to be particularly beneficial for students least likely to succeed (in college overall or within specific majors, such as STEM). In our study, we refer to these students as "at risk" and report an average ESE for these students (d = .26) that is substantial in comparison to the ES found in existing education research; it is .22 standard-deviation units greater than the ES for students not designated "at-risk." This finding is consistent with prior literature emphasizing that social-psychological interventions are most beneficial when they target specific groups of students for specific reasons (Harackiewicz & Priniski, 2018).

This finding is also in line with Sisk et al. (2018), who found evidence that academically high-risk students and economically disadvantaged students may benefit from growth mindset interventions. They do suggest to interpret this result with caution given that few effect sizes contributed to the results and the sample sizes for low SES students were relatively small. Nonetheless, the finding for this particular group is a promising outcome, given that addressing the achievement gap between advantaged and disadvantaged students is a priority. Social-psychological interventions may offer a low-cost approach for doing so.

More effective at highly-selective institutions. Compared to narrative reviews, meta-analytic reviews are often criticized for stripping away context. Indeed, context is very important for generalizability in our study since most students in the United States attend noncompetitive colleges and universities; the factors that affect student achievement and persistence may certainly differ across contexts. In our meta-analysis, we have attempted to capture context. We found that studies conducted at more-selective institutions had larger ESEs than studies conducted at less-selective ones. This finding makes sense when considering the forces that social-psychological interventions are, according to research literature, operating against. A "belonging" intervention will be countering greater feelings of *not* belonging at a highly-selective institution than at a broader-access institution at which a student may have many high school class mates. Similarly, the stereotype threat faced by URM students and women in STEM fields is more prevalent at selective institutions because, in most cases, the proportion of racial minority and low-income students is small (Inzlicht & Ben-Zeev, 2000; Walton, Spencer, & Erman, 2013).

In our study, the less selective institutions include broad access institutions and community colleges. Given that these types of colleges enroll more than 50% of college students, many of whom are often underprepared for college level coursework (Bailey, 2009), in addition to being from low-SES families and often the first in the family to attend college, the finding about selectivity has implications for how we should think about student support (Deil-Amen, 2011). That is, the majority of students in most colleges may need more than these "light touch" interventions to successfully progress through the college to degree pipeline. Indeed, persistence models for community college students stress the importance of strong student-faculty interactions and the need for these interactions to be nurtured *in the classroom* given that most community college students commute to school. High-quality interactions and the experiences they help to create in the classroom are important for community college student success (Deil-Amen, 2011).

Social-psychological interventions may still be appropriate in these contexts, but institutional administrators and practitioners need to consider how to tailor them to fit their students. For example, qualitative studies have discussed the importance of cultivating the development of a "college-going identity" for non-traditional students

(Collatos et al., 2004; Saunders & Serna, 2004). As a result, interventions that are based on identity frameworks could prove to be more impactful.

Further, the finding about selectivity coupled with the low ESE for studies conducted in non-lab settings (d = .11 as compared .23 for studies conducted in a lab setting) has implications regarding generalizability, scalability, and research. Most importantly, this large split in ESE indicates that the expected effects from future social-psychological interventions in non-lab college settings are only about half of our overall estimate. Second, the split provides an important opportunity for research on program design: that is, are there elements of lab delivery that can be applied in more universal delivery, in order to increase the effectiveness of future non-lab interventions? Perhaps of greatest concern, the magnitude of this difference slightly weakens our confidence in the conclusions of the other moderator analyses we conducted. That is, design or participant elements that overlap substantially with laboratory settings could falsely appear to be more effective than their design or participant opposites, simply because of the non-laboratory difference.

Lastly, we must be cognizant of the fact that future growth of social-psychological interventions in college *may* change our conclusions. As we mentioned earlier, Sisk et al. (2018) reported a much lower ESE for growth mindset interventions than Lazowski & Hulleman (2016) does (.08 versus .56).[9] We speculated that this was largely due to the increase in study sample size and questioned whether the ESE for other types of interventions would diminish in a similar fashion in updated syntheses. This seems to be true, although to a lesser degree, for social belonging interventions. Lazowski & Hulleman

[9] See Table 4 for a review of ESEs across key pieces of literature.

(2016) include 3 social belonging studies in the college context; we include 7 and the ESEs differ by .07 (.23 versus .16).

Social-psychological interventions show an average effect that is large enough to make a difference for students. The results of our study are therefore relevant for any institutional administrator who works (directly or indirectly) to help students succeed in terms of development, retention, engagement, achievement, and graduation. A good approach moving forward may be to build on the identified heterogeneity in social-psychological interventions to improve them so that they can be beneficial for a broader group of students.

References

Studies with an asterisk are included in the meta-analysis.

*Acee, T. W., & Weinstein, C. E. (2010). Effects of a value-reappraisal intervention on statistics students' motivation and performance. *Journal of Experimental Education, 78*(4), 487–512. doi:1.1080/00220970903352753

Amrai, K., Motlagh, S. E., Zalani, H. A., & Parhon, H. (2011). The relationship between academic motivation and academic achievement students. *Procedia – Social and Behavioral Sciences, 15*, 399–402. doi:1.1016/j.sbspro.2011.03.111

*Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology, 38*(2), 113–125. doi:1.1006/jesp.2001.1491

Bailey, T. (2009). Challenge and opportunity: Rethinking the role and function of developmental education in community college. *New Directions for Community Colleges,* (145), 11-30.

Baum, S., Cunningham, A., & Tanenbaum, C. (2015). Educational attainment: Understanding the data. *Change: The Magazine of Higher Learning, 47*(6), 18–27. doi: 1.1080/00091383.2015.1089755

Baum, S., & Ma, H. (2007). *Education pays, 2007: The benefits of higher education for individuals and society.* New York, NY: The College Board. Retrieved from http://www.collegeboard.com/prod_downloads/about/news_info/cbsenior/yr2007/ed-pays-2007.pdf

Baum, S., Ma, J., & Payea, K. (2013). *Education pays, 2013: The benefits of higher education for individuals and society.* New York, NY: The College Board. Retrieved from https://trends.collegeboard.org/sites/default/files/education-pays-2013-full-report.pdf

Blackwell, L. S., & Trzesniewski, K. H. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78*(1), 246–263. doi:1.1111/j.1467-8624.2007.00995.x

*Boese, G. D. B., Stewart, T. L., Perry, R. P., & Hamm, J. M. (2013). Assisting failure-prone individuals to navigate achievement transitions using a cognitive motivation treatment (attributional retraining). *Journal of Applied Social Psychology, 43*(9), 1946–1955. doi:1.1111/jasp.12139

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive meta-analysis version 2*. Englewood, NJ: Biostat.

Borenstein, M., Higgins, J. P. T., & Rothstein, H. (2009). *Introduction to meta-analysis*. Chichester, UK: John Wiley & Sons.

*Brady, S. T., Reeves, S. L., Garcia, J., Purdie-Vaughns, V., Cook, J. E., Taborsky-Barba, S., … Cohen, G. L. (2016). The psychology of the affirmed learner: Spontaneous self-affirmation in the face of stress. *Journal of Educational Psychology, 108*(3), 353–373. hdoi:1.1037/edu0000091

*Broda, M., Yun, J., Schneider, B., Yeager, D. S., Walton, G. M., & Diemer, M. (2018). Reducing inequality in academic success for incoming college students: A randomized trial of growth mindset and belonging interventions. *Journal of Research on Educational Effectiveness*. Advance online publication. doi:1.1080/19345747.2018.1429037

*Canning, E. A., Harackiewicz, J. M., Priniski, S. J., Hecht, C. A., Tibbetts, Y., & Hyde, J. S. (2017). Improving performance and retention in introductory biology with a utility-value intervention. *Journal of Educational Psychology*. Advance online publication. doi:1.1037/edu0000244

Center for Collegiate Mental Health (CCMH). (2019, January). *2018 Annual Report* (Publication No. STA 19-180).

Chen, X., & Soldner, M. (2013). *STEM attrition: College students' paths into and out of STEM fields. (NCES 2014-001)*. Washington, DC: National Center for Educational Statistics. Retrieved from https://nces.ed.gov/pubs2014/2014001rev.pdf

Cooper, H. & Rosenthal, R. (1980). Statistical verses traditional procedures for summarizing research findings. *Psychology Bulletin, 87,* 442-449.

Deil-Amen, R.J. (2011). Socio-academic integrative moments: Rethinking academic and social integration among two-year college students in career-related programs. *Journal of Higher Education, 82(*1), 54-91.

*Durik, A. M., Shechter, O. G., Noh, M., Rozek, C. S., & Harackiewicz, J. M. (2015). What if I can't? Success expectancies moderate the effects of utility value information on situational interest and performance. *Motivation and Emotion, 39*(1), 104–118. doi:1.1007/s11031-014-9419-0

Duval, S., & Tweedie, R. (2000a). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95*(449), 89–98. doi:1.1080/01621459.200.10473905

Duval, S., & Tweedie, R. (2000b). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455–463. doi:10.1111/j.0006-341x.2000.00455.x

Dweck, C. S. (2010). *Mindset: The new psychology of success*. New York, NY: Ballantine Books.

Eagan, M.K., Stolzenberg, E.B., Zimmerman, H.B., Aragon, M.C., Whang Sayson, H., Rios-Aguilar, C. (2017). *The American freshman: National norms fall 2016.* Los Angeles: Higher Education Research Institute, UCLA.

Eccles, J. S., Adler, T., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation*. (pp. 75–146). San Francisco, CA: W. H. Freeman.

Egger, M., Smith, G.D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*(7109), 629-634. doi:10.1136/bmj.315.7109.629.

*Eskreis-Winkeler, L., Young, V., Brunwasser, S. M., Shulman, E. P., Tsukayama, E., & Duckworth, A. L. (2016). Using wise interventions to motivate deliberate practice. *Journal of Personality and Social Psychology, 111*(5), 728–744. doi:1.1037/pspp0000074

*Folger, W., Carter, J., & Chase, P. (2004). Supporting first generation college freshmen with small group intervention. *College Student Journal, 38*(3), 472–477.

Gershenfeld, S., Ward Hood, D., & Zhan, M. (2016). The role of first-semester GPA in predicting graduation rates of underrepresented students. *Journal of College Student Retention: Research, Theory, & Practice, 17*(4), 469-488.

*Gripshover, S., Beaubien, J. B., Romero, C. L., Yeager, D. P., Dweck, C. S., Walton, G. M., & Paunesku, D. (2017). *The growing impact of learning mindset interventions among community college students*. Manuscript in Preparation.

Hall, N. C., Hladkyj, S., Perry, R. P., & Ruthig, J. C. (2004). The role of attributional retraining and elaborative learning in college students' academic development. *The Journal of Social Psychology, 144*(6), 591–612. doi:1.3200/SOCP.144.6.591–612

Hall, N. C., Perry, R. P., Chipperfield, J. G., Clifton, R. a., & Haynes, T. L. (2006). Enhancing primary and secondary control in achievement settings through writing–based attributional retraining. *Journal of Social and Clinical Psychology, 25*(4), 361–391. doi:1.1521/jscp.2006.25.4.361

Hamm, J. M., Perry, R. P., Clifton, R. A., Judith, G., & Boese, G. D. (2014). Attributional retraining: A motivation treatment with differential psychosocial and performance benefits for failure prone individuals in competitive achievement settings. *Basic and Applied Social Psychology, 36*(4), 221–237. doi:1.1080/01973533.2014.890623

*Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Giffen, C. J., Blair, S. S., Rouse, D. I., & Hyde, J. S. (2014). Closing the social class achievement gap for first-generation students in undergraduate biology. *Journal of Educational Psychology, 106*(2), 375–389. doi:1.1037/a0034679

*Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Priniski, S. J., & Hyde, J. S. (2015). Closing achievement gaps with a utility-value intervention: Disentangling race and social class. *Journal of Personality and Social Psychology, 111*(5), 745–765. doi:1.1037/pspp0000075

Harackiewicz, J. M., & Priniski, S. J. (2018). Improving student outcomes in higher education: The science of targeted intervention. *Annual Review of Psychology, 69*, 409–435. doi:1.1146/annurev-psych-122216-011725

Harbord, R. M., & Harris, R. J. (2009). Updated tests for small-study effects in meta-analyses. *Stata Journal, 9*(2), 197–21. Retrieved from http://www.stata-journal.com/article.html?article=sbe19_6

Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London, UK: Routledge.

Higgins, E. T., & Rholes, W. S. (1978). “Saying is believing”: Effects of message modification on memory and liking for the person described. *Journal of Experimental Social Psychology, 14*(4), 363–378. doi:1.1016/0022-1031(78)90032-X

Ho, T. K. & Krishna, D. (2016, April 20). *Risk factors associated with first-year college dropout*. Poster session presented at Virginia Commonwealth University Poster Symposium for Undergraduate Research and Creativity, Richmond, VA.

*Hulleman, C. S., & An, B. P. (2017). *The effects of utility value on interest and performance: A causal analysis*. Manuscript in Preparation.

*Hulleman, C. S., Kosovich, J. J., Barron, K. E., & Daniel, D. B. (2016). Making connections: Replicating and extending the utility value intervention in the classroom. *Journal of Educational Psychology,* 109(3), 387–404. doi:1.1037/edu0000146

Hurtado, S., & Carter, D. F. (1997). Effects of college transition and perceptions of the campus racial climate on Latino college students' sense of belonging. *Sociology of Education, 70*(4), 324–345. doi:1.2307/2673270

Inzlicht, M., & Ben-Zeev, T. (2000). A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science, 11*(5), 365-371.

Karabenick, S. A., & Urdan, T. (Eds.) (2014). *Motivational interventions (Advances in motivation and achievement, Volume 18)*. Bingley, UK: Emerald Group Publishing Limited. doi:1.1108/S0749-7423201418

*Layous, K., Davis, E. M., Garcia, J., Purdie-Vaughns, V., Cook, J. E., & Cohen, G. L. (2017). Feeling left out, but affirmed: Protecting against the negative effects of low belonging in college. *Journal of Experimental Social Psychology, 69*, 227–231. doi:1.1016/j.jesp.2016.09.008

Lazowski, R. A., & Hulleman, C. S. (2016). Motivation interventions in education: A meta-analytic review. *Review of Educational Research, 640*, 602–64. doi:1.3102/0034654315617832

Lee, H., Warschauer, M., & Lee, J. H. (2018). The effects of corpus use on second language vocabulary learning: A multilevel meta-analysis. *Applied Linguistics*. Advance online publication. doi:10.1093/applin/amy012

Lipsey, M. & Wilson, D. *Practical meta-analysis*. Thousand Oaks: CA: Sage; 2001.

Liu, O. L., Bridgeman, B., & Adler, R. M. (2012). Measuring learning outcomes in higher education: motivation matters. *Educational Researcher, 41*(9), 352–362. doi:1.3102/0013189X12459679

*Menec, V. H., Perry, R. P., Struthers, C. W., Schonwetter, D. J., Hechter, F. J., & Eichholz, B. L. (1994). Assisting at-risk college students with attributional retraining and effective teaching. *Journal of Applied Social Psychology, 24*(8), 675–701. doi:1.1111/j.1559-1816.1994.tb00607.x

* Miyake, A., Kost-Smith, L. E., Finkelstein, N. D., Pollock, S. J., Cohen, G. L., & Ito, T. A. (2010). Reducing the gender achievement gap in college science: A classroom study of values affirmation. *Science*, *330*(6008), 1234–1237. doi:1.1126/science.1195996

*Murphy, M. C., Carter, E. R., Gopalan, M., Walton, G. M., & Bottoms, B. (2017). *Testing the effects of a social belonging intervention on persistence and achievement in a broad access university context*. Manuscript in Preparation.

National Academies of Sciences, Engineering, and Medicine (NASEM) report. (2017). *Supporting students' college success: The role of assessment of intrapersonal and interpersonal competencies*. Washington, DC: The National Academies Press. doi:1.17226/24697.

National Research Council. (2012). *Education for life and work: Developing transferable knowledge and skills in the 21st Century*. Washington, DC: The National Academies Press. doi:1.17226/13398

Niu, L., Behar-Horenstein, L. S., & Garvan, C. W. (2013). Do instructional interventions influence college students' critical thinking skills? A meta-analysis. *Educational Research Review, 9*, 114–128. doi:1.1016/j.edurev.2012.12.002

Paunesku, D. (2013). Scaled-up social psychology: Intervening wisely and broadly in education. (Doctoral dissertation). Retrieved from http://www.stanford.edu/~paunesku/paunesku_2013.pdf

*Perry, R. P., & Magnusson, J.-L. (1989). Causal attributions and perceived performance: Consequences for college students' achievement and perceived control in different instructional conditions. *Journal of Educational Psychology, 81*(2), 164–172. doi:1.1037/0022-0663.81.2.164

*Perry, R. P., Stupnisky, R. H., Hall, N. C., Chipperfield, J. G., & Weiner, B. (2010). Bad starts and better finishes: Attributional retraining and initial performance in competitive achievement settings. *Journal of Social and Clinical Psychology, 29*(6), 668–70. doi:1.1521/jscp.201.29.6.668

Pratt, T. C., Turanovic, J. J., & Cullen, F. T. (2016). Revisiting the criminological consequences of exposure to fetal testosterone: A meta-analysis of the 2D:4D digit ratio. *Criminology, 54*(4), 587–62. doi:1.1111/1745-9125.12115

Pratt, T. C., Turanovic, J. J., Fox, K. A., & Wright, K. A. (2014). Self-control and victimization: A meta-analysis. *Criminology, 52*(1), 87–116. doi:1.1111/1745-9125.12030

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods (2nd edition).* Newbury Park, CA: Sage.

Robbins, S. B., Allen, J., Casillas, A., Peterson, C. H., & Le, H. (2006). Unraveling the differential effects of motivational and skills, social, and self-management measures from traditional predictors of college outcomes. *Journal of Educational Psychology, 98*(3), 598–616. doi:1.1037/0022-0663.98.3.598

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130*(2), 261–288. doi:1.1037/0033-2909.13.2.261

*Ruthig, J. C., Perry, R. P., Hall, N. C., & Hladkyj, S. (2004). Optimism and attributional retraining: Longitudinal effects on academic achievement, test anxiety, and voluntary course withdrawal in college students. *Journal of Applied Social Psychology, 34*(4), 709–73. doi:1.1111/j.1559-1816.2004.tb02566.x

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78. doi:1.1037/0003-066X.55.1.68

Shadish, R. W., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

Shapiro, D., Dundar, A., Huie, F., Wakhungu, P., Yuan, X., Nathan, A., & Hwang, Y. A. (2017, April). *A National view of student attainment rates by race and ethnicity – Fall 2010 cohort* (Signature Report No. 12b). Herndon, VA: National Student Clearinghouse Research Center. Retrieved from https://nscresearchcenter.org/wp-content/uploads/Signature12-RaceEthnicity.pdf

Sisk, V., Burgoyne, A., Sun, J., Butler, J., & Macnamara, B. (2018). To what extent and under which circumstances are growth mind-sets important to academic achievement? Two meta-analyses. *Psychological Science, (29),* 549-571.

Sneyers, E., & De Witte, K. (2017). Interventions in higher education and their effect on student success: A meta-analysis. *Educational Review, 70*(2), 208–228. doi:1.1080/00131911.2017.1300874

StataCorp. (2015). *Stata statistical software: Release 14*. College Station, TX: StataCorp LP.

*Stephens, N. M., Hamedani, M. G., & Destin, M. (2014). Closing the social class achievement gap: A diversity education intervention improves the academic performance of first-generation college students and the college transition for all students. *Psychological Science, 25*(4), 943–953. doi:1.1177/0956797613518349

Stewart, S., Lim, D.H., & Kim, J. (2015). Factors influencing college persistence for first-time students. *Journal of Developmental Education,* 12-20.

*Struthers, C., & Perry, R. P. (1996). Attributional style, attributional retraining, and inoculation against motivational deficits. *Social Psychology of Education, 1*(2), 171–187. doi:1.1007/BF02334731

Tyson, C. (2014, September 10). To maximize graduation rates, colleges should focus on middle-range students, research shows. *Inside Higher Ed*. Retrieved from https://www.insidehighered.com

*Walton, G. M., & Cohen, G. L. (2007). A question of belonging: Race, social fit, and achievement. *Journal of Personality and Social Psychology, 92*(1), 82–96. doi:1.1037/0022-3514.92.1.82

*Walton, G. M., Logel, C., Peach, J. M., Spencer, S. J., & Zanna, M. P. (2015). Two brief interventions to mitigate a “chilly climate” transform women’s experience, relationships, and achievement in engineering. *Journal of Educational Psychology, 107*(2), 468–485. doi:1.1037/a0037461

Walton, G. M., Spencer, S. J., & Erman, S. (2013). Affirmative meritocracy. *Social Issues and Policy Review, 7*(1). 1-35.

Wigfield, A., Eccles, J. S., Fredricks, J. A., Simpkins, S., Roeser, R. W., & Schiefele, U. (2015). Development of achievement motivation and engagement. *Handbook of Child Psychology and Developmental Science, 3*(16), 1–44. doi:1.1002/9781118963418.childpsy316

*Wilson, T. D., & Linville, P. W. (1982). Improving the academic performance of college freshmen: Attribution therapy revisited. *Journal of Personality and Social Psychology, 42*(2), 367–376. doi:1.1037/0022-3514.42.2.367

*Woolf, K., McManus, I. C., Gill, D., & Dacre, J. (2009). The effect of a brief social intervention on the examination results of UK medical students: A cluster randomised controlled trial. BMC Medical Education, 9(35). doi:1.1186/1472-6920-9-35

Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research, 81*(2), 267–301. doi:1.3102/0034654310378174

*Yeager, D. S., Walton, G. M., Brady, S. T., Akcinar, E. N., Paunesku, D., Keane, L., ... & Gomez, E. M. (2016). Teaching a lay theory before college narrows achievement gaps at scale. *Proceedings of the National Academy of Sciences*, *113*(24), E3341–E3348. doi:1.1073/pnas.1524360113

Tables & Figures

Table 1.1
Mean Effect Size Estimates

	(1) N of ESEs	(2) N of studies	(3) Mean ESE	(4) SE	(5) Joint p-value
Overall Mean Estimate	113	42	0.150***	(0.032)	
Intervention Type					0.345
Utility value	14	8	0.196*	(0.078)	
Sense of belonging	20	7	0.162*	(0.078)	
Growth mindset/AR	55	19	0.178***	(0.046)	
Values affirmation	24	8	0.047	(0.063)	
Outcome Measure					0.082
Competency test	13	---	0.313***	(0.084)	
Course exam	15	---	0.141	(0.078)	
Course grade	20	---	0.126*	(0.057)	
GPA	46	---	0.172***	(0.042)	
Persistence measure	19	---	0.064	(0.053)	
Institutional Setting					0.033
Selective	65	19	0.182***	(0.042)	
Not-selective	15	10	0.014	(0.067)	
Design Elements					
Writing component = No	43	15	0.143**	(0.071)	0.976
Writing component = Yes	70	27	0.116***	(0.024)	
Intervention administered in a lab setting = No	71	27	0.111**	(0.036)	0.074
Intervention administered in a lab setting = Yes	42	15	0.232***	(0.055)	
Subgroup Population					0.000
At risk group in college or within major = No	28	---	0.036	(0.049)	
At risk group in college or within major = Yes	40	---	0.258***	(0.047)	

Publication Status					0.484
Published = No	23	9	0.109***	(0.067)	
Published = Yes	90	33	0.165***	(0.039)	
Publication Date					0.371
Publication date: 1985-2009	23	8	0.214**	(0.081)	
Publication date: 2010-2014	35	10	0.198**	(0.065)	
Publication date: 2015-2019	55	24	0.114**	(0.040)	

Note. Multi-level models were used to retrieve average effect size estimates. Column 3 ESE retrieved using bivariate regression models with one independent variable. ESE reported as predicted means. Asterisks indicate whether the coefficient is significantly different from zero. Joint p-value indicates whether the coefficients are significantly different from one another. * p<.05; ** p<.01; *** p<.001 .

Table 1.2
Institution Selectivity Categories

Acceptance Rate	Graduation-Rate	Selectivity Category
0.05	0.95	1=Selective University
0.30	0.94	1=Selective University
0.41	0.86	1=Selective University
0.53	79.6	2=Moderately Selective
0.53	79.6	2=Moderately Selective
0.66	0.84	2=Moderately Selective
0.66	0.79	2=Moderately Selective
0.70	0.59	2=Moderately Selective
0.72	0.82	2=Moderately Selective
Competitive	0.80	2=Moderately Selective

Note. Each institution was categorized as 1 = selective or 2 = moderately selective if it had both the designated acceptance rate and graduation rate. Institutions accepting the majority of students, such as community colleges, were categorized as 3 = broad access.

Table 1.3
List of Studies Included in Meta-Analysis

Study	Intervention Type	# of Effect Size Estimates	Study Effect Size Estimate	Study Standard Error
Boese et al. (2013)	Attribution retraining	4	0.240	0.306
Hamm et al. (2014)	Attribution retraining	4	0.168	0.224
Perry et al. (1989)	Attribution retraining	4	-0.390	0.337
Perry et al. (2010)	Attribution retraining	9	0.617	0.173
Ruthig et al. (2004)	Attribution retraining	4	0.426	0.210
Struthers & Perry (1996)	Attribution retraining	4	-0.003	0.259
Wilson & Linville (1985)	Attribution retraining	2	0.333	0.275
Aronson et al. (2002)	Growth mindset	2	0.466	0.236
Bostwick & Becker-Blease (2015)	Growth mindset	4	0.234	0.247
Broda et al. Study 1 (2018)	Growth mindset	6	0.051	0.054
Burnette et al. (2019)	Growth mindset	1	0.073	0.090
Eskreis-Winkler et al. Study 3 (2016)	Growth mindset	1	0.381	0.189
Fabert (2014)	Growth mindset	1	0.289	0.124
Gripshover et al. Study 3 (2017)	Growth mindset	2	0.010	0.034
Gripshover et al. Study 7 (2017)	Growth mindset	2	-0.020	0.051
Gripshover et al. Study 9 (2017)	Growth mindset	2	0.000	0.046
PERTS (2017)	Growth mindset	1	0.077	0.064
Sriram (2013)	Growth mindset	1	-0.311	0.197
Wilson Study 2 (2009)	Growth mindset	1	1.479	0.328
Broda et al. Study 2 (2018)	Social belonging	6	0.022	0.054
Murphy et al. (2017)	Social belonging	6	0.142	0.091
Stephens et al. (2014)	Social belonging	2	0.337	0.252
Walton & Cohen (2011)	Social belonging	2	0.082	0.299
Walton & Cohen Study 2 (2007)	Social belonging	2	0.033	0.466
Walton et al. Study 2 (2015)	Social belonging	1	0.360	0.150
Yeager et al. Study 3 (2016)	Social belonging	1	0.250	0.082

Acee & Weinstein (2010)	Utility value	2	0.298	0.311
Canning et al. (2017)	Utility value	1	0.260	0.082
Durik et al. Study 1 (2015)	Utility value	2	0.082	0.268
Durik et al. Study 2 (2015)	Utility value	1	0.459	0.174
Harackiewicz et al. (2015)	Utility value	1	0.143	0.063
Hulleman et al. (2016)	Utility value	2	0.246	0.112
Hulleman et al. (2017)	Utility value	1	0.260	0.295
McPartlan et al. (2019)	Utility value	4	0.003	0.190
Brady et al. (2016)	Values-affirmation	2	0.017	0.157
Gripshover et al. Study 4 (2017)	Values-affirmation	2	-0.130	0.077
Harackiewicz et al. (2014)	Values-affirmation	4	0.017	0.100
Layous et al. (2017)	Values-affirmation	2	0.407	0.207
Miyake et al. (2010)	Values-affirmation	6	-0.067	0.146
Tibbetts et al. (2016)	Values-affirmation	2	0.008	0.101
Walton et al. Study 1 (2015)	Values-affirmation	2	0.020	0.150
Woolf et al. (2009)	Values-affirmation	4	0.115	0.155

Note. McPartlan et al. (2019) study includes first-year pilot study.

Table 1.4
Review of Literature & Comparison of ESEs

	Solanki, Jones, & Lee (working paper)		Lazowski & Hulleman (2016), Review of Educational Research[a]		Harackiewicz & Priniski (2018), Annual Review of Psychology[b]		Sisk, Burgoyne, Sun, Butler, & Macnamara (2018), Psychological Science[c]	
	Meta-analysis		Meta-analysis		Narrative review		Meta-analysis	
	N	ES	N	ES	N	ES	N	ES
Overall ES estimate	42	0.150 (0.032)	10	0.205 (0.067)	15	0.150 (0.057)	---	---
<u>Intervention Type:</u>								
Utility value	8	0.196 (0.078)	1	---	3	---	---	---
Social belonging	7	0.162 (0.078)	3	0.232 (0.078)	5	---	---	---
Growth mindset/AR	19	0.178 (0.046)	5	0.313 (0.112)	2	---	---	---
Growth mindset	12	0.127 [0.054]	1	---	1	---	13	0.080 (0.045)
Values affirmation	8	0.047 (0.063)	1	---	5	---	---	---

Note. N = number of studies in postsecondary context. Standard error in parentheses.

[a] N includes randomized control trial studies. They include 16 studies in the postsecondary context; however, 5 do not report an academic outcome and we can retrieve ESE from one.

[b] Harackiewicz & Priniski (2018) include 21 studies; we were unable to retrieve ESE for 5 of them.

[c] The effect size estimate is not going to be the same as that reported in Sisk et al. for the following reasons: 1. We use the most updated version of the manuscript, see Bostwick (2015) for an example; 2. We use hedges *g*, not cohens *d*; and 3. We include studies not included in their analysis.

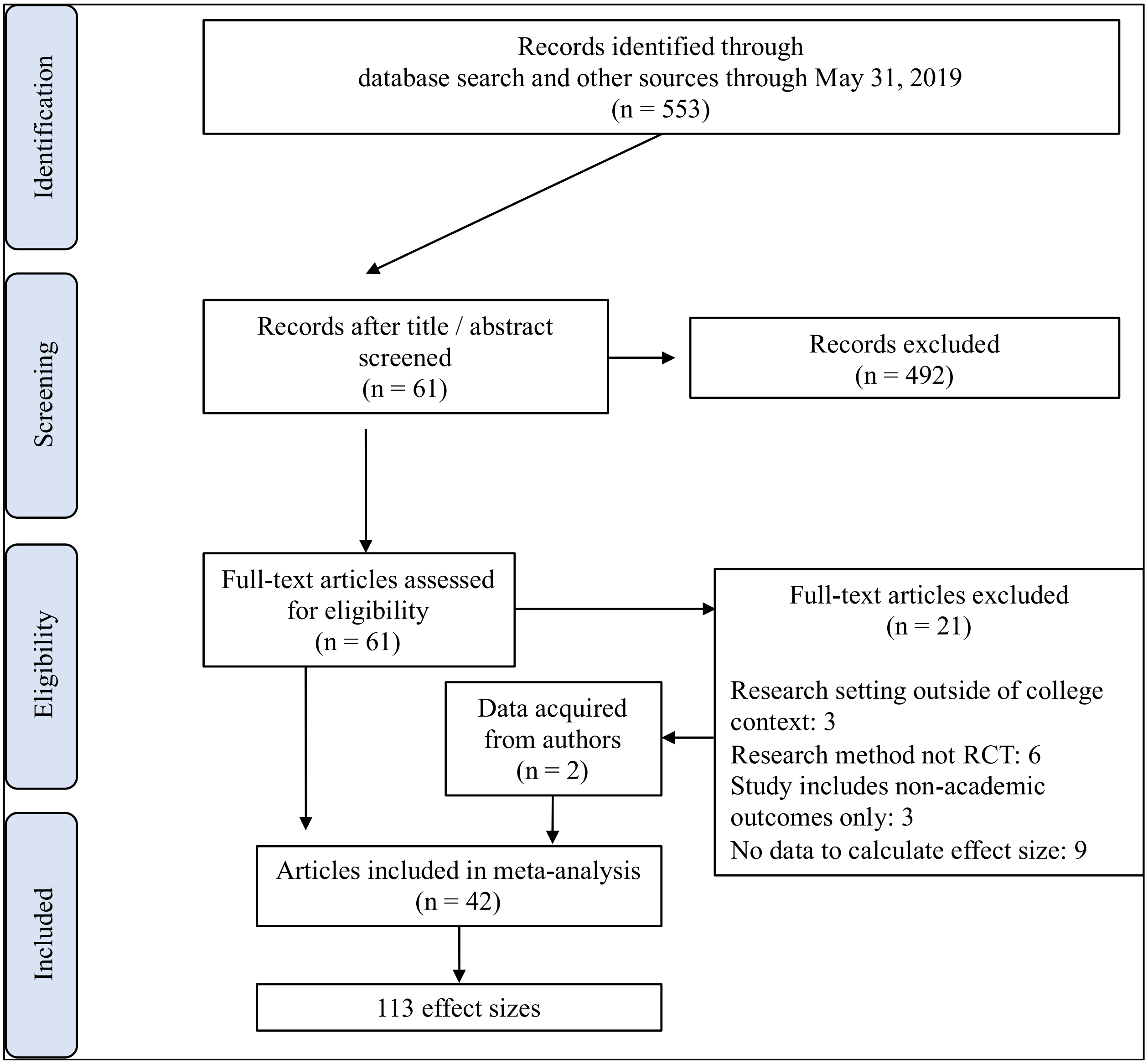

Figure 1.1 PRISMA Flow Chart

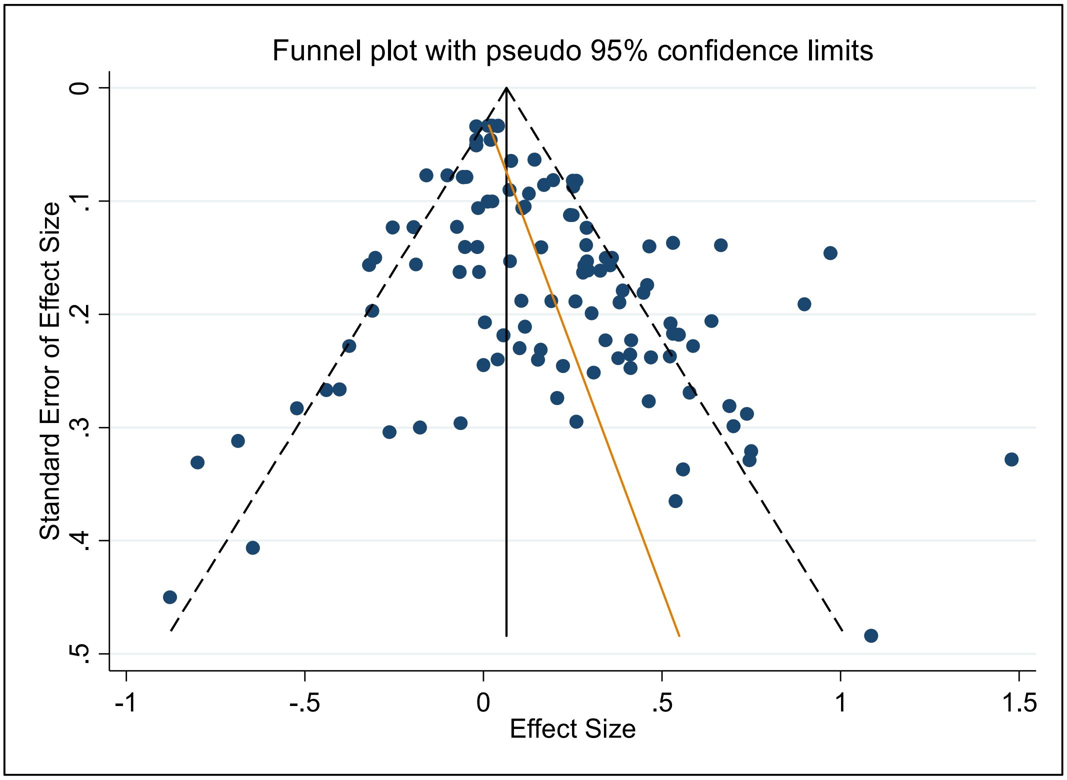

Figure 1.2 Funnel Plot

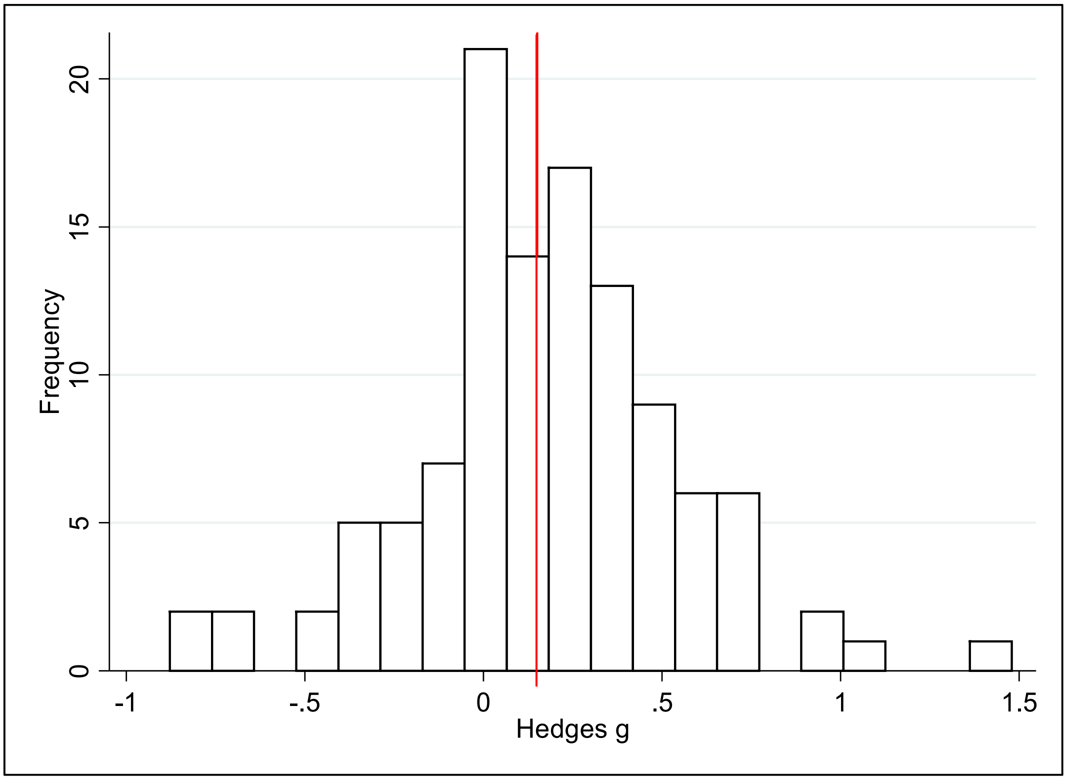

Figure 1.3 Histogram of Effect Size Estimates

CHAPTER 2

Learning Communities in STEM Education:

Evaluating the Impact of the EASE Program on STEM Success

1. Introduction

This fall, approximately 3.8 million American students will attend a four-year college or university intending to major in a STEM field, yet only about half will receive a STEM degree (Chen, 2013; NSF, 2015; NCES, 2015; Seymour & Hewitt, 1997). Completion rates are even lower for under-represented ethnic and racial minority (URM) students and first-generation (FG) students (HERI, 2010; Hurtado, Cuellar, Guillermo-Wann, Velasco; 2010). A number of studies attribute low degree completion rates to students' academic underpreparation as a result of prior educational experiences. Indeed, the American Association of College and Universities reports that 53% of students entering college are academically underprepared, lacking fundamental skills in at least one of the three basic areas of reading, writing, and mathematics (Tritelli, 2003). Developmental education is the common prescription for aiding underprepared students, though whether developmental education helps students' progression through college is still a topic of discussion (Calcagno & Long, 2008; Mattorell & McFarlin, 2011; Xu, 2016).

Along with academic underpreparedness, factors such as personal autonomy, self-confidence, and study behaviors are correlated with course grades and both retention and graduation rates. Also, a number of empirically-based studies have linked sense of belonging to student success in college (e.g., Hurtado & Carter, 1997; Strayhorn, 2012), stressing the importance of altering the college student experience at a program or institution-wide level so students not only receive academic support, but also develop social and intellectual connections with faculty and peers.

The present study evaluates a learning community program designed to improve the educational experience of STEM students at a large public university. Rooted in Tinto's integration theory (1975), the concept of a learning community involves academic and social integration in the university environment; learning communities are designed to foster a strong student support system by promoting student-student and student-instructor interaction through student co-enrollment in two or more courses and involvement in academic and social activities outside the classroom setting. Research relying on case studies and matched-comparison designs has revealed that learning community programs are associated with student academic achievement and retention; further, they improve the overall student experience particularly in the critical first year of college (see Dagley, Georgiopoulos, Reece, & Young, 2015 and Taylor, Moore, MacGregor, & Lindblad, 2003 for a full review).

While these studies are informative, only a small number use research designs that address potential selection bias issues, and only a few were conducted in a STEM context. This study therefore addresses selection issues and provides a rigorous evaluation of a learning community program for the biological sciences major at a large public university in California: the Enhanced Academic Success Experience (EASE) program. EASE is offered to students who fall below a certain SAT Math score threshold (or ACT Math score equivalent), thereby allowing this study's use of a regression discontinuity design (RDD). An RDD, in turn, produces treatment effects that have a causal interpretation. This paper evaluates the impact of EASE on a variety of short- and long-term student outcomes that include academic indicators of success, such as course grades, GPA, retention, and non-cognitive measures, all of which are collected from three waves of survey data. The data includes detailed measures of academic and social integration that have the potential to illuminate the processes by which STEM students benefit from a learning community program such as EASE; they also reveal whether these processes remain influential throughout students' first year of college.

The paper focuses on three research questions:

1. What is the impact of EASE on social-psychological measures of the college experience and short-term measures of academic performance?
2. What is the impact of EASE on long-term measures of academic performance and retention?
3. Do the impacts of the EASE program vary by student subgroup?

2. Theoretical Framework & Relevant Literature

2.1 Learning Communities

Learning communities are an educational strategy designed to improve college student engagement and success. They are inspired by Tinto's integration theory (1975), which links academic and social involvement to academic performance and college student retention (Astin, 1984; Braxton & McClendon, 2001; Carini, Kuh, & Klein, 2006; Tinto, 1975; Tinto & Goodsell, 1993). Students who are more academically and socially involved on campus, particularly with regard to faculty and peers, are more likely to demonstrate greater engagement with learning activities—spending more time studying and putting in more academic effort, for example—a factor that is correlated with academic success and student retention (Barnett, 2011; Engstrom & Tinto, 2007; Kuh, Carini, & Klein, 2004). Academic and social involvement also develops students' sense of belonging in the academic environment. Research has shown that the psychological processes associated with belonging uncertainty have the potential to interfere with academic functioning, as common challenges become much more severe when students feel they are the only ones dealing with them or feel people like them do not often succeed. Low-income, FG, and URM college students and women in male-dominated (i.e., STEM) majors fit this profile (Freeman, Anderman, & Jensen, 2007; Hoffman, Richmond, Morrow, & Salomone, 2003; Stephens, Fryberg, Markus, Johnson, & Covarrubias, 2012). Thus, developing a strong sense of belonging early in one's college career is important, and instituting learning communities may be one way for institutions to communicate that students matter.

Learning communities are designed to increase opportunities for students to interact with peers, faculty, and the curriculum, which allows for construction of a strong support system. The majority of learning communities incorporate active and collaborative learning activities (e.g., students co-enrolling in courses) and promote involvement in complementary academic and social activities that extend beyond the classroom (e.g., students meeting weekly in a study skills course and/or with a group mentor). Faculty involved with learning communities are encouraged to use active pedagogical strategies that foster meaningful interaction between students and instructors. They are also encouraged to engage with one another in considering ways to support student learning outcomes (Smith, Macgregor, Matthews, & Gabelnick, 2004; Zhao & Kuh, 2004).

2.2 Extant Literature about the Impact of Learning Communities on Student Success in College

Prior research has documented that first-year students who participate in learning communities have higher grades, retention rates, and self-reported levels of engagement than peers who do not have a learning community experience. Further, learning community students report studying more with peers outside of class and becoming more involved in academic activities (Engstrom & Tinto, 2007; Shapiro & Levine, 1999; Taylor et al., 2003; Tinto & Goodsell, 1993; Tinto & Russo, 1994; Zhao & Kuh, 2004). While these studies have generally identified a positive correlation between participation in learning communities and student academic/social processes, they use survey data and, further, do not use research designs that support causal inference. Therefore, results may be subject to biases related to students' self-selection into learning communities.

The most rigorous evaluation of learning communities to date comes from a recent report by MDRC (Visher, Weiss, Weissman, Rudd, & Wathington, 2012). Using random assignment, MDRC evaluated the impact of a one-semester learning community program on students assigned to developmental English classes at six different community colleges. In general, the study failed to find any consistent evidence that learning communities positively influenced students' college persistence and academic performance.

A number of plausible reasons could explain these null effects. First, unlike the majority of studies about learning communities, the MDRC study included community college students. Due to the nature of this college context, particular challenges involving student inclusion on campus could be present. In fact, many scholars contend that Tinto's theory about college persistence (1975), which motivates the learning community model, was explicitly designed for the university setting and traditional college-going students. Further, cultivating students' college-going identity and validating the pursuit of college goals is more important in regard to college persistence for community college students than it is for traditional university students (Deil-Amen, 2011; Karp, Hughes, O'Gara, 2010). In general, elements specific to community college students may not have been key components in the MDRC learning community.

Second, students participating in the learning community programs the MDRC study assessed came from a variety of fields. The lack of common interests and goals represented could have substantially weakened the connections between students and sense of belonging, a major component of learning communities that is often cited as being highly correlated with student academic progress and retention decisions (Smith et al., 2004; Strayhorn, 2012). Also, the MDRC report discussed six programs' interventions, all of which lasted only one semester, and programs included only one component of a learning community: paired courses, where students in the learning community were co-enrolled in courses.

It is important to note that regardless of null effects, the MRDC study makes a unique contribution to the literature about learning communities, as it is the only learning community study to use the gold standard in research design: randomization of participants for treatment and control conditions. Findings can therefore be viewed as having a causal interpretation, which is important because this sheds light on the possibility that correlational studies, most of which have shown positive impacts, could be over-stating the benefits students receive from participating in a learning community program. However, as mentioned above, the community college context must be considered when interpreting the results of this study and comparing them to others in the field.

2.3 Potential Value of Learning Communities in STEM Education

Most learning community studies have been implemented for the general population of first-year college students or for those in developmental education programs, such as the MDRC study (2012). Though only a small number of learning community programs have been implemented in STEM programs, the nature of STEM programs makes students well-positioned to reap the benefits of a learning community. This is because students in STEM programs often face discouragement and a loss of confidence due to initially low grades; they experience the weakening of their morale as a result of competitive STEM culture and the generally unwelcoming atmosphere of STEM courses (Fabert, 2014; Hall & Sandler, 1982). Further, students are often overwhelmed by the rigorous curriculum, fast-paced instruction, need for independent work, and content overload in courses taught by sometimes-unengaging STEM faculty (PCAST, 2012; Seymour & Hewitt, 1997).

Discouragement, anxiety, and struggle are often exacerbated for females and students of color who face a particularly challenging climate in STEM classrooms and STEM departments on campus. For example, Hall and Sandler (1982) observed that faculty members generally responded differently to students on the basis of gender. They were less likely to call on female students and ask them challenging questions. As another example, the Herzig (2004) literature review found that female students and students of color who chose to leave doctoral programs in mathematics experienced isolation while in school. They were often left out of social events, had poor relationships with their advisors, and experienced a competitive environment that contradicted their preferred style of interaction. The "chilly climate" phenomenon has been cited as contributing to the relatively low retention rates in STEM fields (Hall & Sandler, 1982). Lastly, female students and students of color also face stereotype threat, which refers to the anxiety one can experience as a member of a stigmatized social group particularly when asked to engage in a task related to the negative stereotype associated with the social group (Steele, 1997). Exposure to stereotype threat is hypothesized to be a reason females and minority students disengage in certain STEM courses and occupations.

Inclusive learning communities provide academic support, headed by faculty interested in effective instruction and strong student-instructor relationships; this indicates that these communities can foster positive student development in STEM. In fact, a substantial body of qualitative literature makes it clear that faculty-student interaction is a key factor in promoting student success particularly among students in most need of support, such as first-generation college students, students of color, and students underprepared for college-level coursework (Fries-Britt & Turner, 2002; Dayton, Gonzalez-Vasquez, Marinez, and Plum, 2004). Faculty also play a role in furthering students' aspirations, including their desire to major in a certain area and commitment to degree completion (Anaya & Cole, 2000; Arredondo, 1995; Braxton, Bray, & Berger, 2000).

Learning communities also foster the development of peer networks, such as study groups, around a common goal and identity, which is important to student success (Sacerdote, 2001; Stadtfeld et al., 2018). For example, students who are oriented towards a common goal often feel more obligated to succeed and help others in their network succeed (Miller, Groccia, & Miller, 2001). Also, strong peer networks foster feelings of belonging and often have access to substantial resources and information. These networks could therefore be particularly beneficial for FG students, many of whom begin their college career requiring additional academic support and are uncertain about how to successfully navigate the college experience (Warburton, Bugarin, & Nunez, 2001).

2.4 Extant Literature about Learning Communities in STEM Education

Extant literature about learning communities in STEM education includes only one study, to my knowledge, that uses a rigorous research design. Russell (2017) evaluated the impact of a freshman learning community program for STEM students called the Experimental Study Group (ESG) using an instrumental variables approach at the Massachusetts Institute of Technology. The program offers multi-dimensional support and includes a number of components from a traditional learning community model, including co-enrollment in courses, effective pedagogy, academic support, and mentorship.

Russell (2017) estimated effects of ESG participation on GPA, total credits completed during college, and degree attainment, finding no statistically significant effects. The author did, however, report significant impacts for subgroup populations of students, such as female students, minority students, and low-income students; students from these groups who participated in ESG earned higher GPAs than otherwise similar students who did not participate in the program. Also, ESG participation increased the probability that a minority student would major in a STEM major associated with higher career earnings—such as computer science or electrical engineering—by 37 percentage points. These estimates support the promise of learning communities—especially for students of color who, at a predominantly white institution, are most likely in greatest need of belonging support and academic integration. However, they must be interpreted with awareness of the small sample size they are derived from.

Unlike Russell (2017), other studies about learning community programs in STEM rely primarily on case studies and matched-comparison designs (Dagley et al., 2015; Taylor et al. 2003). Because the designs for these studies are subject to bias related to students' self-selection into the program, what remains unclear is the ability of learning communities to positively influence STEM academic achievement, STEM persistence, and the overall STEM student experience. Further, at this point, the studies that do use rigorous designs seem to suggest that program impact is largely dependent on program elements and that those who are targeted as most likely candidates to benefit from a program. Further, because the number of studies providing causal estimates is small, there is a lack of agreement about what is known theoretically and what is actually proven to work empirically.

The present study makes a number of unique contributions to existing literature about learning communities in STEM education. First, it adds to the small body of work in this field, outlined above, using a rigorous research design that supports causal inference. Unlike the studies discussed above, the present study takes place in a different context—a four-year public university. Additionally, it is intended to not only illuminate learning community impact by evaluating the EASE program using a number of academic outcomes, but to also examine measures of student social-psychological experience in college that, to date, are absent in STEM learning community literature. Lastly, given that students in the EASE program are all biological sciences majors, this study has the potential to serve as a case study for other institutions considering using the learning community model in STEM fields to improve student outcomes. It therefore is also intended to advance current knowledge about learning communities, exploring whether they are an effective way to improve the higher education STEM experience.

3. Research Setting & Program Description

3.1 Research Setting

This study was conducted at the University of California, Irvine (UCI) School of Biological Sciences. Among STEM majors at UCI, the biological sciences major attracts the largest group of students, is the second largest major on campus, and contains a disproportionate number of URM and FG students. In a typical UCI biological sciences freshman class, roughly 50% of students are FG students, roughly 30% are URM students, roughly 10% are White students, and roughly 40% are low-income students. Historically, the UCI Department of Biological Sciences has struggled with retention. For example, in the past decade, roughly 35% of incoming freshmen designated as majoring in the biological sciences did not graduate with a biological sciences degree after four years of study. Further, during this same period of time, URM students changed their major designation in greater numbers than their White and Asian-American counterparts did. Thus, the School of Biological Sciences at UCI is an ideal setting for investigating issues related to persistence and inequalities in STEM education.

3.2 The EASE Program

The EASE program targets all biological sciences freshmen with an SAT Math score below 600 (or ACT Math score equivalent). This standardized test score is used to determine eligibility for a number of reasons. Most notable among them is that the UCI Department of Biological Sciences found that SAT Math score is correlated with student performance in introductory major courses. Also, in 2015, all UCI campus departments were no longer allowed to use thresholds to determine major eligibility, as prior to 2015, students intending to declare a major in the biological sciences were required to have a minimum SAT Math score of 550. In anticipation of having a cohort of students who might be academically underprepared, the department considered a number of support systems, deciding that a learning community program could be particularly effective. Lastly, the department considered available resources, such as peer mentors and threshold markers, that could result in the maximum number of students the program could feasibly support. A threshold equal to 600 includes roughly 40% of freshman students in the biological science major.

EASE participants are provided multi-dimensional support that includes the following resources[10]:

(1) Academic remediation: EASE students are required to take an additional developmental chemistry course online the summer prior to college matriculation. This course is designed to prepare potential biological sciences majors for college-level courses in chemistry and biology.[11]

(2) Academic and social support: Prior to the start of the school year, the Department of Biological Sciences determines which students are eligible to participate in EASE. They then group these students into cohorts of 30 students. Each cohort is enrolled in the same five biology and chemistry courses (lectures and discussion sections) for one year. Cohorts are also matched with a senior biological sciences mentor. Mentors are upperclassman biological sciences majors selected by the department; they have a tutoring background and have excelled in introductory biological sciences courses. The mentors provide increased academic support and serve as students' main guide to campus resources and opportunities. Lastly, EASE students participate in a weekly 50-minute seminar led by an EASE mentor. Seminar topics are generally academic in

[10] A participant is defined as one who is assigned to the EASE program, enrolled in the weekly Freshman seminar course, and enrolled in both first-year biology and chemistry courses.
[11] Cohort 1 (2016) took Developmental Chemistry its first term.

nature and focus particularly on study skills, metacognition, and research experience. Also discussed are general first-year issues, such as how best to communicate with professors and TAs and how to manage fast-paced coursework in a 10-week quarter.

3.3 Data & Sample Description

Full implementation of the EASE program began in fall 2016 (cohort 1) as a collaborative effort between the UCI School of Biological Sciences and the UCI School of Education. The program was also implemented again in 2017 (cohort 2). This paper estimates the impact of EASE on academic and non-academic outcome measures using a pooled dataset that includes both cohort 1 and cohort 2 (N = 1970).

The UCI Registrar's Office provided student-level data including demographics and standardized test scores as well as UCI transcript data, such as information about students' major. Program-level data detailing students' actual enrollment in the EASE program was also provided. Lastly, records of students' social-psychological measures were collected during both the 2016-2017 and the 2017-2018 academic year. A team of researchers administered three waves of surveys online through the university's course management system (full survey available in Appendix 2.1).[12]

Table 2.1 provides descriptive statistics on demographic characteristics for students in the biological sciences major (N = 1970). As shown in Table 2.1, the biological sciences major is comprised of a diverse group of students; 12% of

[12] The response rate for cohort 1 and cohort 2 was 85% and 80%, respectively, at wave 1.

students are White, 34% are URM, and 54% are Asian. Additionally, 50% of students are first-generation.

3.4 Outcome Measures

Two sets of student outcome measures are analyzed: social-psychological outcomes (also referred to as non-academic outcomes) and academic outcomes.

Social-psychological outcomes measure a variety of student attitudes and behaviors regarding the field of biology. These are scaled measures for which individual items and Cronbach's alphas[13] are listed in Appendix 2.1.

Belonging in biology assesses the extent to which students felt they belonged in the discipline of biology at UCI. Items were adapted from Hoffman et al. (2002)'s Sense of Belonging Scale to ensure they are specific to the biology discipline instead of to the university in general.

Academic integration measures the frequency with which participants engaged in various school-related activities during their first term on campus, such as talking to faculty, planning coursework with academic advisors, and attending study groups.

Academic outcomes consist of performance and persistence data, including the following measures:

Bio 93 grade is the student's final score for the introductory biology course From DNA to Organisms, which is the first of two introductory biology courses. This

[13] Cronbach's alpha is an index of internal consistency reliability that assesses the degree to which responses are consistent across a set of multiple measures for the same construct. (Warner, 2013).

course took place during the fall quarter for both cohorts, and grades were assessed using a standard 4-point scale.

Bio 94 grade is the student's final score for the second introductory biology course, From Organisms to Ecosystems. This course took place during the winter quarter for both cohorts, and grades for the course were assessed using a standard 4-point scale.

Overall GPA is the student's cumulative GPA at the end of the first academic year. Biological sciences majors typically take chemistry, biology, and mathematics courses their freshman year.

Retention is a dichotomous variable measuring whether students were still intending to declare a major in the biological sciences (coded as 1) or had changed their major (coded as 0). Retention is measured at the end of spring quarter.

4. Methodology

4.1 Regression Discontinuity Design (RDD)

Because the EASE program follows a standardized test cutoff score for assignment, this study utilizes an RDD to estimate the causal impacts of the program regarding student outcomes. In the current research context, if it is assumed that nothing other than EASE assignment varies discontinuously at the cutoff, any observed discontinuity in outcomes at the cutoff may be attributed to the EASE program. For example, while it might be expected that first-year major persistence is positively related to students' standardized test math score, the EASE program is the only plausible explanation for a discontinuous jump in this relationship for students that fall just around the score cutoff. The basic implementation of the RDD identifies the impact of the EASE program by comparing the outcomes of students who score slightly above the cutoff score with those who score slightly below it; these students sharply differ in regard to EASE assignment yet are otherwise very similar. The regression coefficient can be interpreted as the causal impact of the intervention for students on the margin of passing the cutoff score (Levin & Calcagno, 2008).

Following Imbens and Lemieux (2008), I focus primarily on a local linear regression that is limited to a narrow bandwidth around the cutoff to estimate the effect of EASE assignment on outcome measures. The equation used is as follows:

$$Y_i = \alpha + \beta_1 (Below_i) + \beta_2 (ScoreDistance_i) + \beta_3 (ScoreDistance_i * Below_i) + X_i + \mu_i \quad (1)$$

Y is the outcome measure (i.e., course grade, first-year GPA, or first-year persistence); *Below* is a binary indicator of whether or not the student was assigned to the EASE program; *ScoreDistance* is the difference between the student's standardized test score and the EASE cutoff score; the interaction term between EASE assignment (*Below*) and the running variable (*ScoreDistance*) allows different slopes above and below the cutoff score. X_i is a vector of individual-level covariates, listed in Table 2.1. β_1 identifies intent-to-treat (ITT) impacts.

As specified above, a local linear regression within a specified bandwidth is used. The mean square error (MSE) optimal bandwidth selection procedure described by Calonico, Cattaneo, & Titiunik (CCT) (2014) produces an optimal bandwidth that is equal to ± 80. Following standards established by What Works

Clearinghouse (WWC, 2015), I also conducted robustness checks with which I estimate treatment effects at 50% of the optimal bandwidth and 200% of the optimal bandwidth. Separate robustness checks related to model fit were conducted in which second-order polynomial terms for the test score distance variable were added to allow the regression function to be more flexible in capturing possible nonlinear relationships between the running variable (SAT Mat score) and outcome measures.[14]

Lastly, it is important to note that RD design provides estimates of the local average treatment effect (Imbens & Angrist, 1994). This means that the conclusions drawn speak only to the subpopulation of students around the cutoff point.

4.2 Addressing Noncompliance

The traditional RD method, known as a sharp RD, assumes full compliance with recommendations based on the test cutoff. In the context of the current study, however, not all students below the cutoff math score enrolled in EASE after being assigned to the program.[15] Accordingly, the average probability of enrollment in EASE is less than one below the cutoff. To deal with potential bias associated with noncompliance, this study followed existing literature that uses a "fuzzy RD" design (Imbens & Lemieux, 2008), utilizing EASE assignment as an instrumental variable for actual participation in EASE and employing a two-stage least squares strategy to estimate the impact of EASE on student outcomes:

---

[14] Akaike's information criterion (AIC) goodness of fit measure was used to determine which model fit the data better.

[15] Noncompliant students (approximately 17% of students assigned to EASE) were those who were able to bypass EASE despite a standardized math score below the cutoff. Reasons for this waiver included earning a 3 or higher on the AP Chemistry test, earning a 700 or higher on the SAT Chemistry Subject Exam, or successfully completing a course equivalent to Chemistry 1P prior to the fall quarter.

$$\mathrm{Enroll_i} = \gamma_0 + \gamma_1(\mathrm{Below_i}) + \gamma_2(\mathrm{ScoreDistance_i}) + \gamma_3(\mathrm{Below_i} * \mathrm{ScoreDistance_i}) + \mathrm{X_i} + \mu_i \qquad (2)$$

$$\mathrm{Y_i} = \delta_0 + \delta_1(\widehat{\mathrm{Enroll_i}}) + \delta_2(\mathrm{ScoreDistance_i}) + \delta_3(\mathrm{Below_i} * \mathrm{ScoreDistance_i}) + \mathrm{X_i} + \varepsilon_i \qquad (3)$$

Enroll indicates enrollment in EASE. Equation (2) represents the first stage in which a linear probability model was used to predict EASE enrollment as a function of EASE assignment. Equation (3) then estimated the local average treatment effect of the predicted probability of enrollment on student outcome measures in the second stage. δ_1 captures the impact of participating in EASE on student outcomes.

4.3 Validity Checks

Three tests were conducted to ascertain the validity of the RDD. First, it was verified that the probability of EASE program enrollment is discontinuous at the SAT Math score passing cutoff; a discontinuity means that SAT Math score is a strong assignment mechanism determining treatment assignment. Figure 2.1 plots the likelihood of participating in the EASE program as a function of SAT Math score (centered to be zero at the passing cutoff). The graph clearly shows a discontinuity at the passing cutoff for EASE enrollment. In fact, the correlations between recommendation and participation are strong, with an overall compliance rate of 83%.

Second, distributions of baseline covariates as a function of SAT Math score were examined, similarly to what would be done in a traditional balance test. The corresponding regressions were also run. As shown in Figure 2.2, the majority of graphical illustrations show a continuous relationship, indicating no statistical differences between students in close proximity to either side of the cutoff score. Table 2.2 provides statistical estimates that support what is found graphically. As shown in Table 2.2, *female* is the only variable that is statistically significant.

Lastly, the density of data around the cutoff was examined to determine whether participants manipulated the assignment score. It is highly unlikely that students had any knowledge of the EASE program or its eligibility requirement at the time they took the standardized exam of their choice, and thus they would have had no incentive to study for the SAT or ACT with the goal of ensuring their Math score would fall on one side of the treatment cutoff. Nonetheless, to be thorough, I determined whether there is a discontinuity in the density of observations at the cutoff by conducting a manipulation test following Cattaneo, Jansson, & Ma (2017). The p-value equal to 0.07 indicates a smooth distribution (see Figures 2.3 and 2.4).

4.4 Treatment Effect Heterogeneity

RDD approach. We examine treatment effect heterogeneity using the RDD approach. This analysis explores whether gaps in social-psychological measures and academic achievement are wider or narrower for certain student subgroups (i.e., female students, URM students, first-generation college students, and low-income students) from participation in the EASE program. Specifically, for this analysis, we include a given individual attribute (such as 'female') in equation 1.

$$\begin{aligned} Y_i = \alpha &+ \beta_1 (Below_i) + \beta_2 (ScoreDistance_i) + \beta_3 (Attribute_i) + \\ &\beta_4 (Below_i * ScoreDistance_i) + \beta_5 (Below_i * Attribute_i) + \\ &\beta_6 (ScoreDistance_i * Attribute_i) + B_7 (Below_i * ScoreDistance_i * Attribute_i) \\ &+ X_i + \mu_i \end{aligned} \tag{4}$$

The β_7 coefficient is the coefficient of interest and indicates whether EASE reduced the academic achievement gap (or in the case of gender, the gender achievement gap).

Difference-in-difference approach. A limitation to estimating treatment effect heterogeneity using the RDD approach described above is that estimates for subgroups are retrieved from small sample sizes. As a robustness check, we use a difference-in-difference approach (DID), thereby increasing the sample size used for the heterogeneity analysis.

DID is a quasi-experimental design that makes use of longitudinal data from treatment and control groups over time. In the present study, the DID analysis examines the impact of the EASE program by looking at changes in both the control group (students with greater than 600 on SAT-math) and treatment group (students with less than 600 on SAT-math) over time, that is, during years when EASE was not offered (2012-2014) and during years when EASE was offered (2016 and 2017).[16] The equation takes the following form:

$$Y_i = \alpha + \beta_1 (EASEELG_i) + \beta_2 (Post_i) + \beta_3 (Attribute_i) + \beta_4 (EASEELG_i * Post_i) + \beta_5 (EASEELG_i * Attribute_i) + \beta_6 (Post_i * Attribute_i) + B_7 (EASEELG_i * Post_i * Attribute_i) + X_i + \mu_i$$

where *EASEELG* is a student-level indicator of eligibility to the EASE program; *Post* denotes time period and is equal to 1 for the Fall 2016 and 2017 cohort and 0 for all the cohorts before 2015, when EASE was not offered. The coefficient, β_7, then indicates whether the performance gap between EASE eligible and ineligible

[16] We omit the 2015 academic year from the analysis. In 2015, the Biological Sciences department implemented the EASE program using volunteer participants.

students for subgroup populations of students (Attribute) was mitigated due to the EASE program. A limitation to this analysis is that I only have academic-related data, specifically, course grades and retention.

Validity checks. Valid estimates in a DID design depend on the suitability of the chosen control group. In the present study, this means students in the same cohorts who earned SAT-math scores equal to or greater than 600. We show in Figure 2.5—top left panel—that a sharp jump in EASE participation is only observed for students with less than a 600 SAT-math score (i.e., the EASE-eligible students). Figure 2.5 provides evidence that we have an appropriate control group.

In addition, it is important for the treatment and control groups to exhibit parallel trends prior to the implementation of the EASE program. That is, in the absence of treatment, the difference between the treatment and control group should be constant over time. To test the parallel trends assumption, I plotted the proportion of each student pre-treatment variable (SAT-math score, gender, race) over time for both ease-eligible students and non-eligible students using the full sample of students (cohorts 2012-2017, excluding 2015). Of the student pre-treatment variables, only one exhibited a parallel trend. For this reason, I restrict the analysis to include students within the optimal bandwidth as specified under the RDD approach (± 80), assuming these students are similar. I test the parallel trends assumption within the specified bandwidth and the results are presented in Figure 2.5. Parallel trends can be seen for all but one pre-treatment variable, white.

5. Results

Tables 2.3-2.5 show results for the short- and long-run non-academic and academic outcome measures. In all tables, columns 3 and 4 include point estimates for analyses that include the sample within the optimal BW, which is equal to ± 80 points. Columns 1 and 2 include point estimates for analyses for which I use a sample that lies within 50% of the optimal BW, and columns 5 and 6 include point estimates for which I use a sample of students that lies within 200% of the optimal BW. For each BW, I ran a model with first-order SAT Math polynomial terms (columns 1, 3, and 5) and second-order SAT Math polynomial terms (columns 2, 4, and 6).

5.1 Non-Academic Outcome Measures

As mentioned earlier, responses to pre- and post-surveys during the fall quarter were used to examine the impact of EASE on two measures of students' experience in college: *sense of belonging* and *academic integration*. Figure 2.6 plots these outcomes as a function of SAT score, providing visual evidence of discontinuities at the cutoff. As shown by the graphs, there is wide variation among individuals in terms of response. I find, however, that there are clear discontinuities at the cutoff score, indicating that students assigned to EASE benefitted from the program.

The statistical estimates confirm what can be seen graphically. As shown in Table 2.3, students assigned to the EASE program reported values for the measure *sense of belonging* that are 0.38 standard deviation units higher (Column 3) than the values reported by those not assigned to the program. This point estimate remains consistent across all model specifications. Similarly, students reported values 0.36 standard deviation units higher for the measure of *academic integration* (Panel B, Column 3); however, the point estimates are less consistent across model specifications. At a BW equal to 40, the results for *academic integration* are much smaller and no longer statistically significant.

5.2 Academic Outcome Measures

Figure 2.6 also illustrates the relationship between EASE participation and student academic outcomes. As shown in Figure 2.6, a small discontinuity exists at the cutoff for Biology 93, students' first introductory biology course in the biological sciences major track. A much more pronounced discontinuity is present for Biology 94, students' second course in the major track. The statistical estimates in Table 2.4 confirm what is visible graphically: the point estimate for Biology 94 (Panel B, Column 3) is equal to 0.28 units on a 0 to 4 point scale (equivalent to moving a student from a B- to a B, for example).

In regard to cumulative year 1 GPA and retention, as shown in Table 2.5, I find that students in EASE earned higher GPAs than students not in the program (Panel A, Column 3). Lastly, I find no impact on retention (Panel B).

5.3 Differential Treatment Effects by Cohort

Thus far, all analyses have used the pooled sample of students. Prior research, however, notes the importance of examining cohort differences. Analyzing heterogeneity by cohort can shed light on whether program impacts may improve over time. Indeed, after the first year of implementation, program administrators can learn from past mistakes and successes and become more efficient in their processes. They can also learn how to make learning community programs more effective. Thus, it is not uncommon to see program impacts increase after the first year of implementation. On the other hand, researchers have speculated that programs are most robust during their first year when administrators and those involved are most excited about them. Examining cohort differences has the potential to shed light on the direction of both EASE and similar programs.

Table 2.6 includes point estimates for our non-academic and academic outcome measures for cohort 1 (column 1) and cohort 2 (columns 2 and 3) using an optimal bandwidth equal to 60 and 80, respectively.[17] As shown in Table 2.6, treatment effects are driven by cohort 1 across all academic outcome measures. For the measure *sense of belonging*, treatment effects are significant across both cohorts but the point estimate is half the magnitude for cohort 2.

5.4 Long-term Academic Outcome Measures

The relationship between EASE and long-term academic outcomes was also examined. These outcomes include cumulative year 2 GPA and student retention in the biological sciences major at the end of students' second year in college. The statistical estimates are reported in Table 2.7. It is important to note that only students in cohort 1 are included here because cohort 2 recently finished its second year in college and data is not yet available.

The impact of EASE on GPA remains at the end of year 2, and the magnitude of the effect is fairly similar to end of year 1 results. Specifically, for cohort 1, students in EASE earned cumulative year 1 GPAs that were 0.33 units higher (Table

---

[17] For cohort 2, the optimal bandwidth using CCT produces a bandwidth equal to 80. Table 2.6 also includes results for cohort 2 using an optimal bandwidth equal to 60 (Column 3).

2.6, Column 1) than those of students not in the program (on a 0-4 point scale). At the end of year 2, the point estimate is equal to 0.32 units. There is no impact on year 2 retention.

5.5 Treatment Effect Heterogeneity

As discussed earlier, one purpose of this study is to better understand whether certain student subgroups (e.g., URM students, female students, and first-generation students) benefit particularly from the learning community experience; the results are presented in Table 2.8. As mentioned earlier, the β_7 coefficient (presented in Table 2.8) is the coefficient of interest and indicates whether EASE reduced the gender achievement gap, for example. The results are null except for one distinction: the EASE program differentially impacts URM students *sense of belonging*.

Table 2.9 presents results using the DID approach.[18] The results are consistent in terms of sign and statistical significance. The magnitude of point estimates, however, are larger across all outcomes.

6. Discussion & Conclusion

The study discussed in this paper advances current knowledge regarding learning community programs with a rigorous evaluation of a first-year STEM learning community program. It investigates outcomes beyond traditional academic outcomes reported in prior studies in the field and includes direct measures of the student experience in college.

[18] The data is repeated cross-section. I have the same group of students observed in multiple time periods. Thus, standard errors are clustered at the individual level to account for within-person correlated error terms.

Findings suggest that the EASE program does indeed improve the student experience in college. Students in the EASE program report higher values for the measure of *sense of belonging* and *academic integration*. Although there is wide variation in terms of student responses, it is most interesting that the EASE program appears to weaken the negative association between student social-psychological measures and SAT Math score, the latter of which is often a proxy for academic preparation. With the sense of belonging measure, for example, there is a clear steep upward slope on the right side of the cutoff score where non-EASE students lie (Figure 2.6). In the absence of treatment, the regression line would likely continue this trend and we would see very low values for sense of belonging for students on the left side of the cutoff. As Figure 2.6 shows, however, this is not the case. Instead, for students on the left side of the cutoff—that is, EASE students—the correlation between students' sense of belonging and SAT Math score is relatively weak.

In addition to the findings outlined above, I find that the EASE program impacted academic outcomes, affecting introductory first-year biology courses and cumulative GPA, for example. It did not, however, have an impact on retention. This result may not be surprising given that only 8% of students left the biological sciences after year 1. By the end of year 2, 20% of students had left the major. Yet, students were equally likely to leave the major, regardless of intervention group. Intervention research has noted that studies are more likely to detect effects on outcomes that are proximal to the intervention (Kraft, 2018; Ruiz-Primo, Shavelson, Hamilton, & Klein, 2002). And the EASE effect could fade over time especially since the program is offered during students' first-year as biological sciences majors.

Regardless, this outcome highlights the complex relationship between academic and social integration and student retention decisions outlined in Tinto's integration model (1975). Future research could benefit from in-depth focus group interviews with EASE students—those that stayed in biological sciences and those that left—to better understand the role of EASE in regard to student retention decisions.

Differences by cohort. The observed cohort differences deserve further attention. As mentioned in section 5.3, the main results are primarily driven by cohort 1. In speculating about why outcomes are markedly different for cohort 2, I first considered differences in program implementation that might have been present. One notable difference between cohort 1 and cohort 2 is that students in cohort 1 took a remedial chemistry course during its first term, whereas cohort 2 did not. Cohort 2 enrolled in Chemistry 1A in the fall, which is the first course of the chemistry sequence. One possibility is that cohort 1 had a relatively easier mix of courses during its first term, which allowed students to focus more on their biology courses and perform better in them. Data on future cohorts of EASE students, in addition to class-level data, might be able to help us better understand these cohort differences.

Treatment effect heterogeneity. For the most part, I did not find any differential treatment effects by subgroup populations of students, specifically by female students or URM students. Social-psychological research offers a potential explanation. Specifically, prior research suggests that when numerical representation is adequate, the psychological pressures that belonging, for example, act upon are less likely to depress students' achievement (Inzlicht & Ben-Zeev, 2000; Logel, Walton, Peach, Spencer, & Zanna, 2010). The belonging intervention described in Walton, Logel, Peach, Spencer, and Zanna (2015) substantiates this hypothesis. They found the GPA of women in *male-dominated* majors rose, on average, more than a full letter grade higher than the GPA of otherwise similar women in the control group. In addition to the effect on GPA, the intervention improved women's perceived experience of their engineering major—as measured by a sense of belonging, self-efficacy, and enjoyment—relative to the control condition. They found null effects across outcomes for females in *gender-diverse* majors.

Thus, the over-representation of females in the biological sciences at UCI, where they represent 70% of students, could be a plausible reason why I did not detect any differences in outcomes by gender. Whether the same reasoning applies to URM students is less clear. URM students represent roughly 35% of biological sciences majors. Is this number an adequate representation? The differential increase in *sense of belonging* (Table 2.8)—albeit extremely small—suggests otherwise. Prior research has found that treatment effects unfold over time (Murphy & Zirkel, 2015; Walton & Cohen, 2011). As such, the social-psychological benefits experienced by students traditionally underrepresented in STEM, such as URM students, may translate to positive long-run academic performance outcomes, such as strong grades earned in the second year of college or even major persistence.

Overall, the findings from this study are informative for researchers, policymakers, and higher education administrators in a number of ways. With the exception of the computer/information sciences field, attrition rates are quite similar across STEM disciplines. For example, nationally, 45% of entering freshmen intending to study the biological sciences leave postsecondary education (15%) or switch to another major (30%). The attrition rates are 18% and 28%, respectively, for matriculating freshmen intending to major in the physical sciences (NCES, 2013). EASE therefore serves as a potential case study for researchers and practitioners in terms of promoting student success in STEM. Other disciplines may want to consider forms of holistic support, of which learning communities are a small part, when considering best ways to serve students and improve persistence within STEM fields.

References

Calcagno, J. C., & Long, B. T. (2008). The impact of postsecondary developmental coursework using a regression discontinuity approach: Addressing endogenous sorting and noncompliance. Mimeo.

Cattaneo, M. D., M. Jansson, & X. Ma. (2017). Simple Local Polynomial Density Estimators. Working paper, University of Michigan.

Chen, X. (2013). STEM Attrition: College Students' Paths Into and Out of STEM Fields. Washington, DC. Retrieved from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2014001rev

Center for Institutional Data Exchange and Analysis (C-IDEA). (2000). *1999-2000 SMET Retention Report*. Norman, OK: University of Oklahoma.

Dadgar, M. (2012). Essays on the economics of community college students' academic and labor market success (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Accession Order No. 3506175)

Darling E., Molina K., Sanders M., Lee F., & Zhao Y. (2008). Belonging and achieving: The role of identity integration. *Social Psychological Perspectives, Advances in Motivation and Achievement*, *15*, 241-273.

Dweck, C. S., & Henderson, V. L. (1988). Theories of intelligence: Background and measures. Unpublished manuscript.

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behavior. In J. T. Spence (Ed.), Achievement and achievement motives: Psychological and sociological approaches (pp. 75–146). San Francisco, CA: W.H. Freeman.

Eccles et al. in document

Fabert, N. (2014). Growth mindset training to increase women's self-efficacy in science and engineering: A randomized-controlled trial. (Doctoral dissertation). Retrieved from https://repository.asu.edu/attachments/137308/content/Fabert_asu_0010E_14151.pdf

Freeman, T. M., Anderman, L. H., & Jensen, J. M. (2007). Sense of belonging in college freshmen at the classroom and campus levels. *The Journal of Experimental Education*, *75*(753), 203–220. http://doi.org/10.3200/JEXE.75.3.203-220

Hall, R. & Sandler, B. (1982). The classroom climate: A chilly one for women? *Project on the Status of Education of Women.* Washington, D.C.: Association of American Colleges.

Herzig, A. H., (2004). Becoming mathematicians: Women and students of color choosing and leaving doctoral mathematics. *Review of Educational Research, 74*(2), 171-214.

Higher Education Research Institute (HERI). (2010). *Degrees of Success: Bachelor's Degree Completion Rates among Initial STEM Majors.* Los Angeles: CA: University of California, Los Angeles.

Hodara, M. (2012). Language minority students at community college: How do developmental education and English as a second language affect their educational outcomes? (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Accession Order No. 3505981)

Huang, G., Taddese, N., & Walter, E. (2000). Entry and persistence of women and minorities in college science and engineering education (No. NCES 2000601). Washington, D.C.: National Center for Education Statistics.

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Priniski, S. J., & Hyde, J. S. (2015, November 2). Closing achievement gaps with a utility-value intervention: Disentangling race and social class. *Journal of Personality and Social Psychology*. Advance online publication.

Hoffman, M., Richmond, J., Morrow, J., & Salomone, K. (2003). Investigating "sense of belonging" in first-year college students. *Journal of College Student Retention*, *4*(3), 227–257.

Huang, G., Taddese, N., & Walter, E. (2000). Entry and persistence of women and minorities in college science and engineering education (No. NCES 2000601). Washington, D.C.: National Center for Education Statistics.

Hurtado S. & Carter D. F. (1997). Effects of college transition and perceptions of the campus racial climate on latino college students' sense of belonging. *Sociology of Education, 70*(4):324-345.

Imbens, G. W., & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrics, 61*(2), 467–476.

Imbens, G. W. & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics, 142*(2), 615-635.

Lee, D. (2008). Randomized experiments from non-random selection in US house elections. *Journal of Econometrics, 142*, 675-697.

Lee, D., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature, 48*, 281–355

Lesik, S. (2007). Do developmental mathematics programs have a causal impact on student retention? An application of discrete-time survival and regression-discontinuity analysis. *Research in Higher Education, 48*(5), 566–591.

Levin, H. & Calcagno, J. C. (2008). Developmental coursework in the community college: An evaluator's perspective. *Community College Review, 35*, 181-207.

Martorell, P., & McFarlin, I. (2011). Help or hindrance? The effects of college remediation on academic and labor market outcomes. *Review of Economics and Statistics, 93*, 436–454.

Master, A., Cheryan, S., & Meltzoff, A. N. (2016). Computing whether she belongs: Stereotypes undermine girls' interest and sense of belonging in computer science. *Journal of Educational Psychology, 108*(3), 424–437. doi.org/10.1037/edu0000061

Miller, J. E., Groccia, J. E., & Miller, M. S. (2001). Introduction. In J. E. Miller, J. E. Groccia, & M. S. Miller (Eds.), *Student-assisted teaching: A guide to faculty-student teamwork.* Bolton, MA: Anker.

Murphy, M. C., & Zirkel, S. (2015). Race and belonging in school: How anticipated and experienced belonging affect choice, persistence, and performance. *Teachers College Record, 117*(12), 1–40.

NCES (2013). Digest of Education Statistics, 2013. Table 306.40. U.S. Department of Education.Washington, D.C.: Institute of Education Sciences

Ostrove, J. M., & Long, S. M. (2007). Social class and belonging: Implications for college adjustment. *The Review of Higher Education, 30*(4), 363–389. http://doi.org/10.1353/rhe.2007.0028

Pascarella, E. T., & Terenzini, P. T. (1980). Predicting freshman persistence voluntary dropout decisions from a theoretical model. *The Journal of Higher Education, 51*(1), 60–75.

President's Council of Advisors on Science and Technology (PCAST). (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics*. Washington, DC: Author.

Robinson, S. (1996). *Underprepared students*. (ED433876) Washington, DC: Educational Resources Information Center.

Sacerdote, B. (2001). Peer effects with random assignment: Results for Dartmouth roommates. *The Quarterly Journal of Economics, 116*(2), 681-704.

Scott-Clayton, J., & Rodriguez, O. (2015). Development, discouragement, or diversion? New evidence on the effects of college remediation Policy. *Education Finance and Policy, 10*, 4-45.

Seymour, E., & Hewitt, N. M. (1997). *Taking about leaving: Why undergraduates leave the sciences*. Boulder, CO: Westview Press.

Sherman, D. K., Bunyan, D. P., Creswell, J. D., & Jaremka, L. M. (2009). Psychological vulnerability and stress: The effects of self-affirmation on sympathetic nervous system responses to naturalistic stressors. *Health Psychology, 28*(5), 554–562.

Silverman A. & Cohen G. (2014). Fostering Positive Narratives: Social-Psychological Interventions to Maximize Motivation in the Classroom and Beyond. *Motivational Interventions*, pp 177-211.

Stadtfeld, C., Voros, A., Elmer, T., Boda, Z. & Raabe, I. J. (2018). Integration in emerging social networks explains academic failure and success. *Proceedings of the National Academy of Sciences, 116*(3), 792-797.

Stephens, N. M., Fryberg, S. A., Markus, H. R., Johnson, C. S., & Covarrubias R. (2012). Unseen disadvantage: how American universities' focus on independence undermines the academic performance of first-generation college students. *Journal of personality and social psychology* 102(6):1178.

Strayhorn, T. L. (2012). *College students' sense of belonging: A key to educational success for all students*. New York: Routledge.

Tinto, V. (1998) Colleges as communities: Taking research on student persistence seriously. *The review of higher education, 21*(2), 167-177.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45*(1), 89–125. http://doi.org/10.3102/00346543045001089

Tinto, V. (2006). Enhancing student persistence: Lessons learned in the United States. *Aná Psicológica, 24*(1), 7-13.

University of California Office of the President (2016, April4). *University of California admits significantly more California freshman students, makes gains in diversity.* Retrieved January 21, 2017, www.universityofcalifornia.edu/press-room/

Walton, G. M. & Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science, 331*(6023), 1447-1451.

Xu, (2016). Assistance or obstacle? The impact of different levels of English developmental education on underprepared students in community colleges *Educational Researcher, 45*, 496–507. doi.org/10.3102/0013189X16683401

Zhao C. M. & Kuh G. D. (2004). Adding value: Learning communities and student engagement. *Research in Higher Education, 45*(2), 115-138.

Zumbrunn, S., McKim, C., Buhs, E., & Hawley, L. R. (2014). Support, belonging, motivation, and engagement in the college classroom: A mixed method study. *Instructional Science, 42*(5), 661–684. doi.org/10.1007/s11251-014-9310-0

Tables & Figures

Table 2.1
Summary Statistics: 2016 & 2017 Biological Sciences Majors

		Cohort	
	Pooled Sample	2016	2017
Female	0.692	0.684	0.699
White	0.122	0.125	0.119
URM	0.338	0.337	0.340
Asian	0.540	0.539	0.541
First-gen status	0.498	0.484	0.510
Low-income status	0.372	0.407	0.341
SAT-math score	602.1	598.0	605.5
N	1970	907	1063

Note. Sample includes undergraduate first-year students, not missing SAT-math.

Table 2.2
Covariate Balance Check

	Female	White	URM	Asian	FG-status	LI status	SB-wave 1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
EASE-eligible	0.133**	0.013	0.034	-0.034	0.018	0.006	0.038
	(0.057)	(0.040)	(0.057)	(0.057)	(0.059)	(0.060)	(0.146)
N	1111	1111	1111	1111	1111	1111	1068

Note. In each model, the demographic characteristic is the outcome variable. Equation (1) is used to estimate the EASE-eligible coefficient within the optimal bandwidth equal to 80 point. FG = First-generation; LI = Low-income; SB = Sense of belonging.

Table 2.3
RD Results: Social-psychological Measures of the Student Experience

	BW = 40		**Optimal BW = 80**		BW = 160	
	(1)	(2)	(3)	(4)	(5)	(6)
	Poly-1	Poly-2	Poly-1	Poly-2	Poly-1	Poly-2
Panel A. Outcome = Sense of Belonging						
ITT	0.418***	1.182***	0.384***	0.394**	0.337***	0.367***
	(0.158)	(0.365)	(0.102)	(0.171)	(0.079)	(0.112)
2SLS	0.567**	2.117**	0.499***	0.524**	0.433***	0.485***
	(0.222)	(1.013)	(0.136)	(0.235)	(0.103)	(0.152)
N	522	522	1019	1019	1535	1535
Panel B. Outcome = Academic Integration						
	BW = 40		**Optimal BW = 80**		BW = 160	
	(1)	(2)	(3)	(4)	(5)	(6)
ITT	0.164	-0.957*	0.364***	0.219	0.378***	0.331**
	(0.193)	(0.549)	(0.127)	(0.213)	(0.097)	(0.140)
2SLS	0.232	-1.857	0.482***	0.306	0.488***	0.443**
	(0.271)	(1.314)	(0.168)	(0.296)	(0.125)	(0.187)
N	541	541	1051	1051	1577	1577

Note. The optimal bandwidth is equal to 80 points. The column headings denote different SAT-math polynomials controlled for in these regressions (all interacted with a dummy for being above/below the threshold). All models include the full list of covariates. Wave 1 sense of belonging measure included as a covariate in Panel A. Robust standard errors used. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table 2.4
RD Results: First-Year Biology Courses

	BW = 40		**Optimal BW = 80**		BW = 160	
	(1)	(2)	(3)	(4)	(5)	(6)
	Poly-1	Poly-2	Poly-1	Poly-2	Poly-1	Poly-2
Panel A. Biology 93 (first-term)						
ITT	0.008	-0.061	0.199*	-0.017	0.160*	0.199
	(0.178)	(0.440)	(0.118)	(0.194)	(0.089)	(0.129)
2SLS	0.011	-0.100	0.257*	-0.022	0.204*	0.261
	(0.238)	(0.717)	(0.153)	(0.256)	(0.114)	(0.171)
Sample Mean			2.94			
N	572	572	1104	1104	1662	1662
Panel B. Biology 94 (second-term)						
ITT	0.151	0.024	0.279**	0.138	0.210**	0.262**
	(0.189)	(0.497)	(0.122)	(0.206)	(0.091)	(0.134)
2SLS	0.196	0.037	0.351**	0.174	0.263**	0.334*
	(0.243)	(0.757)	(0.153)	(0.258)	(0.114)	(0.170)
Sample Mean			2.72			
N	543	543	1042	1042	1562	1562

Note. The optimal bandwidth is equal to 80. The column headings denote different SAT-math polynomials controlled for in these regressions (all interacted with a dummy for being above/below the threshold). All models include the full list of covariates. Robust standard errors used. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table 2.5
RD Results: End of Year 1 Academic Outcomes

	BW = 40		**Optimal BW = 80**		BW = 160	
	(1)	(2)	**(3)**	**(4)**	(5)	(6)
	Poly-1	Poly-2	Poly-1	Poly-2	Poly-1	Poly-2
<u>Panel A. Cumulative year 1 GPA</u>						
ITT	0.109	0.104	0.186**	0.118	0.182***	0.179**
	(0.116)	(0.286)	(0.074)	(0.126)	(0.056)	(0.082)
2SLS	0.153	0.215	0.245**	0.161	0.235***	0.239**
	(0.162)	(0.597)	(0.098)	(0.173)	(0.073)	(0.110)
Sample Mean			2.95			
N	541	541	1033	1033	1543	1543
<u>Panel B. Retention in major, end of year 1</u>						
ITT	-0.002	-0.015	-0.005	-0.010	-0.010	-0.017
	(0.045)	(0.123)	(0.029)	(0.050)	(0.023)	(0.033)
2SLS	-0.003	-0.021	-0.006	-0.013	-0.012	-0.022
	(0.058)	(0.177)	(0.037)	(0.064)	(0.028)	(0.042)
Sample Mean			0.92			
N	541	541	1033	1033	1543	1543

Note. The optimal bandwidth is equal to 80. Retention measured at the end of spring quarter (year 1). The column headings denote different SAT-math polynomials controlled for in these regressions (all interacted with a dummy for being above/below the threshold). All models include the full list of covariates. Robust standard errors used.
* p < 0.10; ** p < 0.05; *** p < 0.01.

Table 2.6
RD Results: Outcome Measures--Differences by Cohort

	Cohort 1 (BW = 60) (1)	Cohort 2 (BW = 80) (2)	Cohort 2 (BW = 60) (3)
Sense of Belonging	0.557***	0.265**	0.220
	(0.195)	(0.134)	(0.163)
Academic Integration	0.325	0.330*	0.333
	(0.227)	(0.173)	(0.205)
Biology 93 (first-term)	0.346**	0.168	-0.247
	(0.174)	(0.165)	(0.213)
Biology 94 (second-term)	0.449**	0.156	0.124
	(0.202)	(0.174)	(0.213)
Cumulative GPA-Year 1	0.331***	0.069	0.057
	(0.123)	(0.100)	(0.125)
Year 1 Retention	0.065**	-0.033	-0.066
	(0.029)	(0.049)	(0.062)
N	387	590	397

Note. The column headings indicate whether the results include cohort 1 or cohort 2. The optimal bandwidth for cohort 1 is equal to 60; the optimal bandwidth for cohort 2 is equal to 80. All results include first-order polynomial terms for SAT-math, in addition to SAT-math interacted with a dummy for being above/below the threshold. I also include the following covariates: a gender dummy variable, race dummy variables, a dummy variable for low-income status, a dummy variable for first-generation status, and wave 1 survey measure (for non-cognitive measures). All estimates are ITT estimates. All models include the full list of covariates. Robust standard errors used. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table 2.7
RD Results: End of Year 2 Academic Outcomes

	BW = 50	**Optimal BW = 60**	BW = 120
	(1)	**(2)**	(3)
Panel A. Cumulative year 2 GPA			
	Poly-1	Poly-1	Poly-1
ITT	0.303**	0.317***	0.195**
	(0.135)	(0.118)	(0.082)
2SLS	0.496**	0.533**	0.318**
	(0.232)	(0.209)	(0.136)
Sample Mean		2.96	
N	262	312	544
Panel B. Retention in major, end of year 2			
ITT	0.017	-0.021	-0.031
	(0.100)	(0.087)	(0.061)
2SLS	0.034	-0.042	-0.053
	(0.195)	(0.169)	(0.103)
Sample Mean		0.80	
N	322	385	699

Note. The analyses use data from cohort 1. The optimal bandwidth is equal to 60. Cumulative year 2 GPA includes students retained in the biological sciences major. Retention measured at the end of year 2. The column headings denote different SAT-math polynomials controlled for in these regressions (all interacted with a dummy for being above/below the threshold). All models include the full list of covariates. Robust standard errors used. * p < 0.10; ** p < 0.05; *** p < 0.01.

Table 2.8
RD Results: Treatment Effect Heterogeneity

	(1)	(2)	(3)	(4)	(5)	(6)
	SB	ASI	Biology 93	Biology 94	Cumgpa-Year 1	Retained
B_7 Coefficient:						
EASE-eligible*SAT-math*Female	0.001	-0.003	-0.008	-0.003	-0.000	0.000
	(0.006)	(0.008)	(0.007)	(0.008)	(0.005)	(0.002)
EASE-eligible*SAT-math*URM	0.016***	0.010	0.003	0.002	0.001	-0.002
	(0.006)	(0.007)	(0.007)	(0.007)	(0.004)	(0.002)
EASE-eligible*SAT-math*Firstgen	0.004	-0.002	-0.005	0.005	0.004	-0.002
	(0.006)	(0.007)	(0.006)	(0.006)	(0.004)	(0.002)
EASE-eligible*SAT-math*Low-income	0.005	-0.007	-0.002	-0.002	0.001	-0.003
	(0.006)	(0.007)	(0.006)	(0.006)	(0.004)	(0.002)
N	1019	1051	1104	1042	1033	1109

Note. The optimal bandwidth is equal to 80 points. The column headings denote different outcome measures. SAT-math is centered at the cutoff score. SB = sense of belonging; ASI = academic and social integration. Wave 1 sense of belonging measure included as a covariate in column 5. Robust standard errors used. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table 2.9
DID Results: Treatment Effect Heterogeneity Analysis

	Biology 93	Biology 94	Retained
Difference-in-difference coefficient			
(EASE-eligible*Treat period)	0.120*	0.216***	-0.005
	(0.069)	(0.069)	(0.023)
N	3013	2700	3149
Triple difference coefficient (B_7)			
(EASE-eligible*Treat period*Female)	0.092	0.114	0.077
	(0.146)	(0.145)	(0.049)
(EASE-eligible*Treat period*URM)	0.022	0.084	-0.058
	(0.152)	(0.153)	(0.051)

Note. Each cell indicates a separate regression. Optimal bandwidth is equal to 80. All models includes the full set of covariates. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

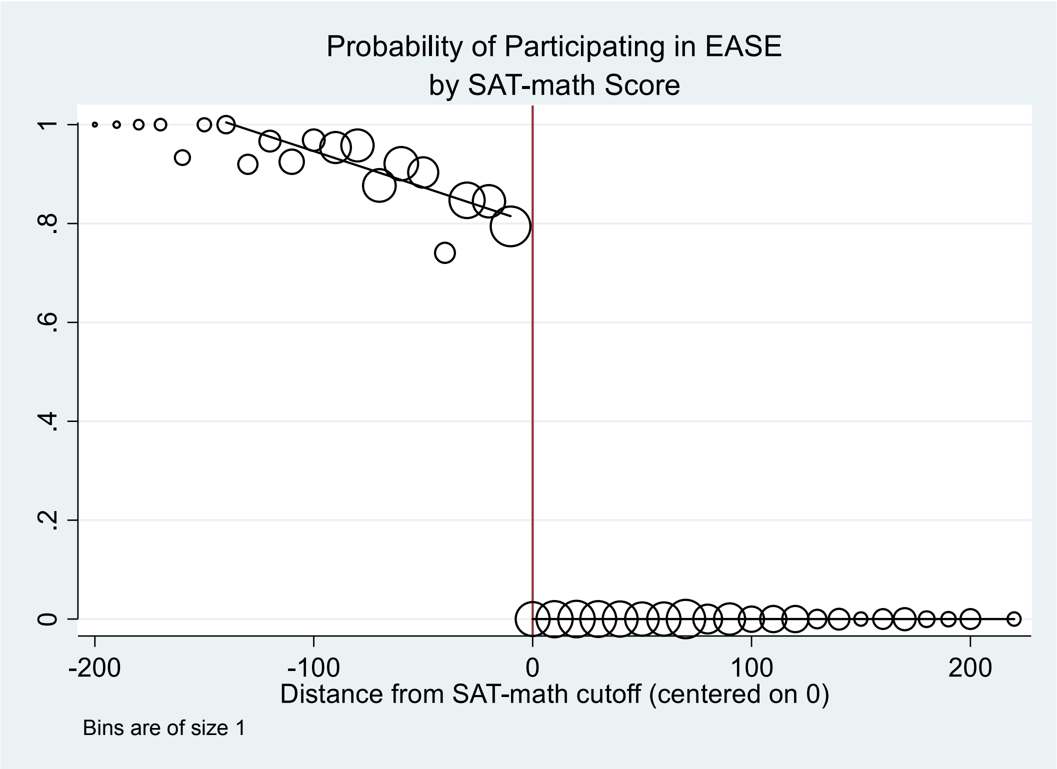

Figure 2.1 Probability of EASE participation by SAT math test score

Figure 2.2 RD validity check: SAT math score distribution by pretreatment individual characteristics

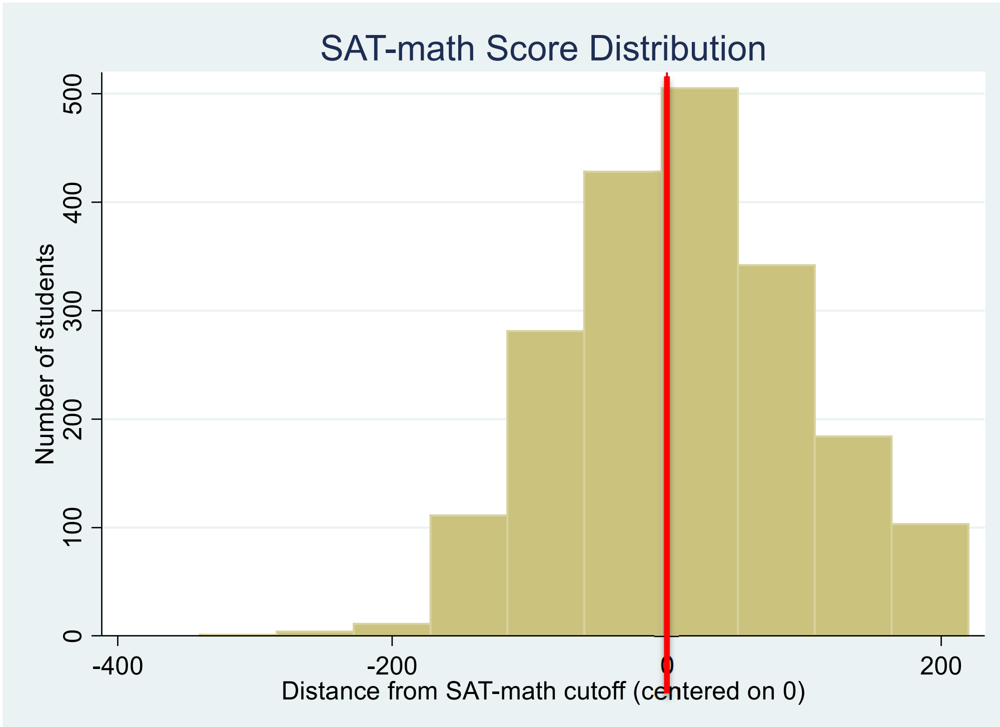

Figure 2.3 RD validity check: Density of SAT math score distribution around the cutoff score

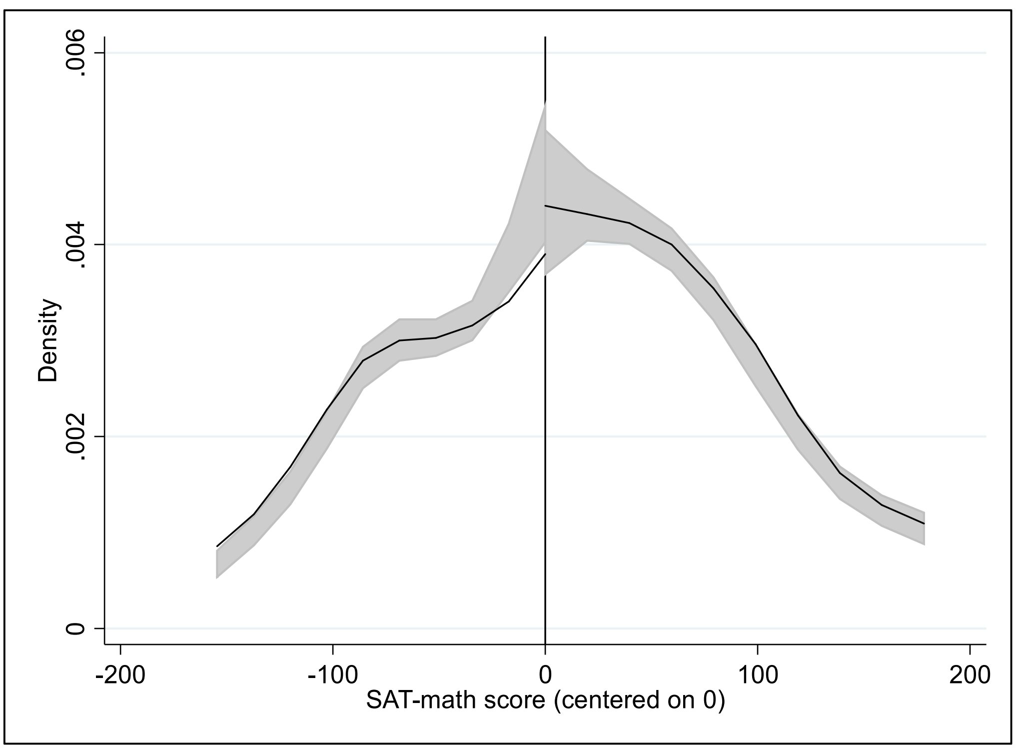

Figure 2.4 Density test (Cattaneo et al., 2017)

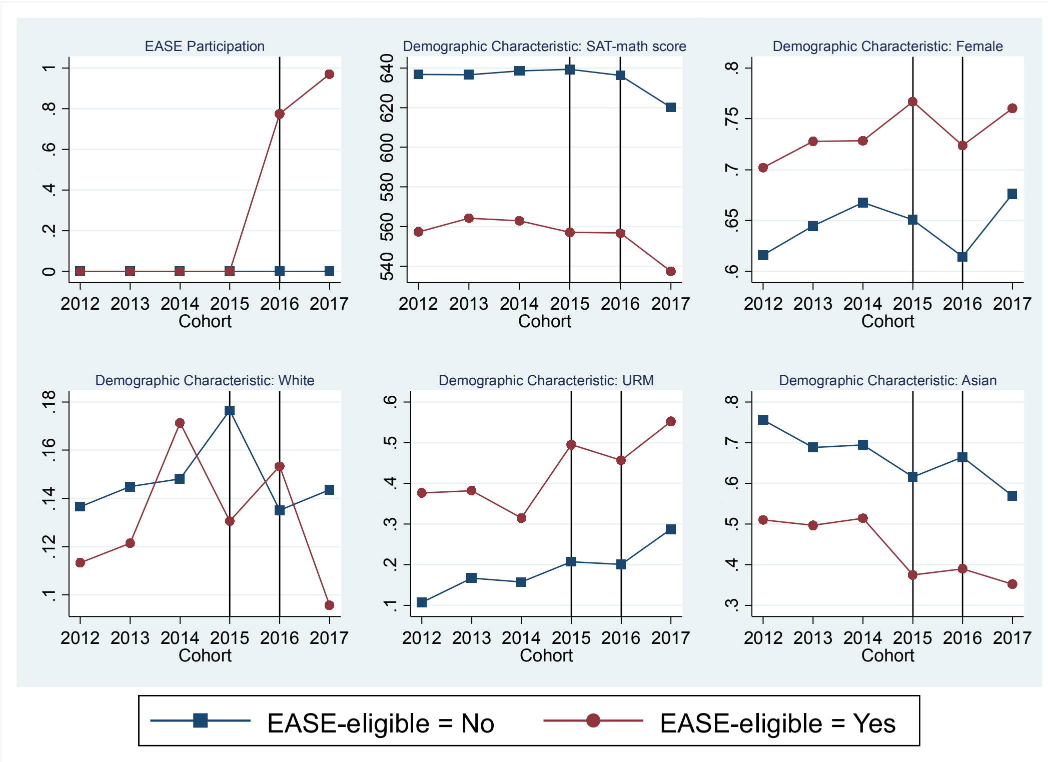

Figure 2.5 Trends in EASE participation and cohort composition before and after implementation of the EASE program

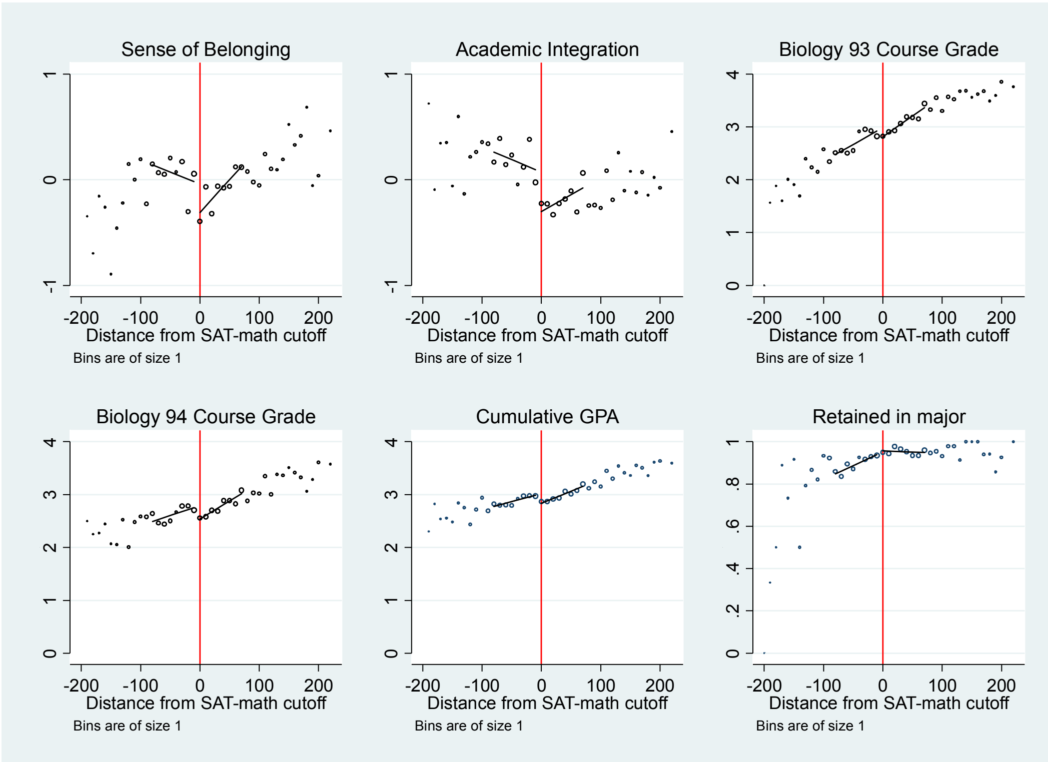

Figure 2.6 Outcome measures by SAT math score

Appendix 2.1: End-of Term Survey by Construct

Motivation - Interest Value (α=0.92)
Instructions: Please indicate how true each statement is for you.
(1. Not at all true ... 7. Very true)

1. I'm really looking forward to learning more about Biology
2. Biology fascinates me
3. I think the field of Biology is very interesting

Sense of Belonging in Biology (α=0.85)
Instructions: Please indicate how true each statement is for you.
(1. Not at all true ... 7. Very true)

1. If I miss a Bio Sci class, I know students who I could get the notes from
2. I discuss events which happen outside of class with my Bio Sci classmates
3. I have developed personal relationships with other students in my Bio Sci classes
4. I feel comfortable volunteering ideas or opinions in my Bio Sci classes
5. I feel comfortable asking a question in my Bio Sci classes
6. I feel comfortable seeking help from my Bio Sci teachers before or after class
7. I feel comfortable asking my Bio Sci teachers for help if I do not understand course-related material
8. I feel comfortable asking my Bio Sci teachers for help with a personal problem

Academic and Social Concerns (α=0.74)
Instructions: Please indicate how true each statement is for you.
(1. Not at all true ... 7. Very true)

1. In college, I sometimes worry that people will dislike me
2. In college, I worry that people will think I'm unintelligent if I do poorly
3. In college, I often get nervous and worried when I talk to people

Growth Mindset of Intelligence (α=0.90)
Instructions: Please indicate how much you agree with the following statements (1=Strongly disagree, 2=Disagree, 3=Mostly disagree, 4=Mostly agree, 5=Agree, 6=Strongly agree)

1. You have a certain amount of intelligence, and you really can't do much to change it (reverse coded)
2. Your intelligence is something about you that you can't change very much (reverse coded)
3. You can learn new things, but you can't really change your basic intelligence (reverse coded)

Academic & Social Integration (α=0.60)
Instructions: Please indicate how often you did the following activities this quarter. Consider all of your classes and activities, not just those for this course.
(0=Never, 1=Once a month, 2=Twice a month, 3=Every week)

1. Talk with faculty about academic matters, outside of class time (including e-mail)
2. Meet with an academic advisor concerning academic plans
3. Meet with a student mentor concerning academic plans (Antleader, peer mentor, EASE, etc.)
4. Attend study groups outside of the classroom
5. Have informal or social contacts with faculty members outside of classrooms and offices

CHAPTER 3

Tenure-Track Appointment for Teaching-Oriented Faculty:

The Impact of Teaching and Research Faculty on Student Outcomes

1. Introduction

While it is widely acknowledged that faculty at research universities assume multiple tasks, including research, teaching, and service, the nature of the relationship between research scholarship and instructional effectiveness is not agreed upon. On one end of this controversial debate is the academic ideal that research and teaching are complementary and hence mutually support each other. For example, it is purported that excellent academic scholars also have the potential to become excellent teachers, grounding their instruction on advanced research-led knowledge (e.g., Brew, 1999, 2013; Brew & Jewell, 2012; Hajdarpasic, Brew, & Popenici, 2015; Halliwell, 2008; Robertson & Bond, 2001; Teichler, Arimoto, & Cummings, 2013). On the other hand, some challenge the belief that research enhances teaching (e.g., Coate, Barnett, & Williams, 2001; Figlio & Schapiro, 2017; Hattie & Marsh, 1996; uz Zaman, 2004). Central to this perspective is the point that the attributes of an outstanding researcher may be distinct from the attributes of an effective teacher, particularly in the context of entry-level courses that aim to create a strong foundation in a discipline. Moreover, since the tenure and promotion criteria often place stronger emphasis on research excellence than teaching excellence at research-focused institutions, research faculty may lack both the time and motivation needed to improve their instructional quality (see uz Zaman, 2004 for a full review of existing evidence on the relationship between teaching and research excellence).

While the question of whether traditional research faculty are indeed in the best position to deliver high quality instruction remains unsettled, colleges are faced with increasing student enrollments. Since tenure-track research faculty often have fixed and relatively low teaching loads at research institutions, colleges have increasingly relied on contingent teaching faculty in the past three decades to address the growing demand for higher education (NCES, 2016). The rapid growth of contingent faculty has created a teaching-focused workforce that now comprise almost three quarters of instructional faculty in the higher education sector (Kezar & Maxey, 2013). However, while contingent faculty provide a flexible and cost-saving way to address the growing population of freshman enrollees, critics have raised concerns regarding their impact on student academic outcomes. These concerns often revolve around the employment conditions of contingent faculty—relatively low compensation, minimal benefits, exclusion from departmental decisions, and lack of job security, few of which are optimally suited for engagement with students and the institution, and may in turn result in unintended negative impact on student learning outcomes and academic progress (Benjamin, 2002, 2003; Bérubé & Ruth, 2015; Schuetz, 2002; Umbach, 2007).

The ongoing conversation about the opportunities and challenges of employing contingent teaching faculty in teaching undergraduate-level courses has led to discussions about whether colleges should consider providing tenure appointments to teaching faculty as a way to professionalize and stabilize the teaching workforce in higher education. For example, in their 2015 book *The Humanities, Higher Education, and Academic Freedom*, Michael Bérubé and Jennifer Ruth explicitly advocate for the formation of a separate tenure track for faculty hired to prioritize teaching instead of research. This faculty group would be in a unique position, as it would be responsible for a significant teaching load similar to that of contingent teaching faculty while being entitled to similar level of job security and academic freedom as tenure-track research faculty.

The University of California (UC) system is one of the pioneers in adopting this dual-tenure-track system. Within the UC system, instructional faculty can be divided into three broad categories: (i) non-tenure track contingent faculty (referred to as "lecturers" hereafter), (ii) tenure-track/tenured research faculty (referred to as "research faculty" hereafter), and (iii) tenure-track/tenured teaching faculty (referred to as "teaching faculty" hereafter). The teaching faculty series was first implemented in the UC system decades ago; however, a policy change in 2002 solidified its specific role on campus, and this particular faculty series has expanded substantially in the last two decades. In 2016, the UC system had 274 faculty in the teaching faculty title series across ten campuses, with roughly half of faculty employed at two of the ten UC campuses.

Teaching faculty in the UC system are members of the academic senate and are entitled to the same rights and privileges as research faculty. Their primary responsibilities are teaching and administration; however, they are also expected to participate in scholarship activity (UCOP, 2002). Surveys administered to teaching faculty show that the majority focus their scholarship efforts on discipline-based education research (DBER) to identify evidence-based teaching practices (Harlow & Sato, 2017). The combination of specialization in teaching and DBER thus distinguishes teaching faculty from both traditional research faculty and lecturers.

Despite increasingly diversified faculty composition in higher education and the expectation that the new model of tenure-track teaching faculty will serve as a stable teaching force to deliver high-quality instruction, little is known regarding their teaching effectiveness relative to research faculty and contingent lecturers. To date, literature in the field has focused on identifying the nature of faculty work among teaching faculty and better understanding faculty attitudes about their work. These studies typically involve teaching faculty in STEM departments nationwide and teaching faculty in research-intensive Canadian institutions. In both contexts, teaching faculty are growing. In fact, in Canada, teaching faculty are the fastest growing category of faculty positions (Rawn & Fox, 2018).

This paper addresses the gap in the literature by examining whether students learn similarly well from research faculty, tenure-track teaching faculty, and contingent lecturers at one of the UC campuses. To our knowledge, this is the first analysis of the impact of tenure-track teaching faculty relative to traditional research faculty and contingent lecturers on students' academic outcomes. Our analysis therefore not only has direct implications for university hiring policies and faculty structure, but also has the potential to serve as an important first step toward identifying potential ways to improve undergraduate program instruction and quality.

We use detailed student transcript and administrative data that includes six cohorts of students initially enrolled at the institution between 2008 and 2013 and examine the impact of taking one's initial course in a field of study with different types of instructors on students' current course performance, subsequent enrollment and performance in the same field, and longer-term outcomes including total credits earned and major declaration in the same field. The major methodological challenge in estimating the impact of instructor type is student self-selection into courses and the non-random assignment of instructors to different classes. To minimize selection bias, we build on prior studies in the field and use a three-way fixed effects model that controls for student-level fixed effects, initial-course fixed effects, and next-class fixed effects, therefore eliminating unobservable biases that are constant at the individual level (e.g., academic capacity), at the initial-course level (e.g., course difficulty), or at the next-class level (e.g., selection into less difficult classes after initial exposure to a field) (see Figlio et al., 2015; Ran and Xu, in press; Xu, 2019).

Our results suggest that faculty type is associated with student learning outcomes. Specifically, students on the margin of passing an introductory course are more likely to fail if the course is taught by a tenure-track teaching faculty, compared with similar students who take an introductory course with either research faculty or lecturers. Aside from this distinction, the distribution of course grades across instructor type is similar. In terms of subsequent outcomes, students who take their initial course with tenure-track teaching faculty, as compared with either research faculty or lecturers, do better in the second course in the same field and eventually accumulate more credits in that field, although the point estimates are small and only marginally significant.

Our paper is related to a small but growing body of literature that focuses on the impact of different types of instructors on student academic outcomes at the postsecondary level (Bettinger & Long, 2010; Carrell & West, 2010; Figlio, Schapiro & Soter, 2015; Hoffmann and Oreopoulos, 2009; Ran & Xu, in press; Xu, 2019). Our study builds on the existing literature on instructor effectiveness and extends it in two important ways. First, we include in our analysis a new category of faculty that has never been examined before—teaching-oriented faculty with tenure appointment. Given a longstanding debate about whether institutions should grant tenure to teaching-faculty, knowledge about their impact on student academic outcomes relative to other types of faculty provides an empirical foundation for better informed staffing decisions by policy makers and college administrators.

In addition, the mixed and, at times, contradictory findings across existing studies regarding the relative impact of contingent faculty highlight the possibility that there might be substantial variation in the contractual form, working conditions, characteristics, and quality of contingent teaching faculty in different institutional contexts. For example, Ran & Xu (in press) defined adjuncts as those hired in transitory positions; approximately one fifth terminated their employment with the college and never returned to teaching after their first year. In contrast, the majority of adjunct faculty in Figlio et al. (2015) had a long-term relationship with the university. Accordingly, the impact of contingent faculty employed under different contracts and within diverse working conditions might also be distinct from each other in terms of faculty engagement and satisfaction with the institution as well as their impact on student learning outcomes. The field has so far gathered evidence about contingent faculty in elite private institutions (Figlio et al., 2015), open-access and less selective public teaching-focused institutions (Bettinger & Long, 2010; Ran & Xu, in press; Xu, 2019), and a public Canadian institution with a largely Caucasian student and faculty population (Hoffmann and Oreopoulos, 2009). Our study adds to this line of research by estimating the impact of contingent faculty at a selective public research institution.

2. Theoretical Motivation & Extant Literature

2.1 The Role of Faculty in Higher Education

This study draws on theoretical perspectives and relevant research about student learning in higher education institutions. More specifically, it is motivated by a line of inquiry suggesting that faculty play a central role in undergraduate education (Astin, 1993; Kezar & Maxey, 2014; Kuh & Hu, 2001; Pascarella & Terenzini, 1991, 2005). Indeed, similar to the K-12 education, faculty and instructor quality is arguably one of the most important factors that influence student success at the college level, especially given that 40% of students entering college are underprepared for college-level coursework (Carrell & West, 2010; Braxton et al., 2000; Condon, Iverson, Manduca, Rutz, & Willett, 2016; Crockett, 2015; Gordon, 2012; Gyurko MacCormack, Bless, & Jodl, 2016; Oolbekkink-Marchand, Van Driel, & Verloop, 2014; Seymour & Hewitt, 1997; Wang, Pascarella, Nelson Laird, & Ribera, 2015; Loes & Pascarella, 2015; Tinto, 2004; Umbach & Wawrzynski, 2005).

Yet, scholars have noted that college faculty vary substantially in how they deliver course content and interact with students, resulting in the widespread concerns about teaching effectiveness (Arum & Roksa, 2011; Bok, 2006). In fact, one of the most commonly reported reasons by students for changing majors or leaving college is unsatisfactory course experiences (Braxton et al., 2000; Seymour & Hewitt, 1997). Researchers and institutional administrators who focus on improving college instruction have proposed several hypotheses about why some instructors do not teach well. Most notable among them is a relative lack of pedagogical training, inadequate resources, and a lack of incentive structures (i.e., those related to promotion and tenure) that prioritize research activities over teaching related activities and reward scholarly excellence more than instructional effectiveness (Gaff, Pruitt-Logan, Sims, & Denecke, 2003; Finkelstein, 1984; uz Zaman, 2004). As a result, faculty often have insufficient incentives to go through systematic pedagogical training and despite the emerging plethora of evidence-based teaching practices that research has shown to be positively linked to student learning, retention, and graduation—such as approaches involving active and collaborative learning techniques, engaging students in learning experiences, and fostering interactions between students—they are rarely adopted in most of the college classrooms (Crockett, 2015; Easton & Guskey, 1983; Kezar & Maxey, 2014; Pascarella & Terenzini, 1995, 2005; Raisman, 2013; Umbach & Wawrzynski, 2005).

One approach that has been used by many institutions and has received increasing popularity is to hire teaching-focused faculty whose primary responsibility is instruction instead of research. In the following sections, we provide a brief overview of two faculty types most relevant to the current study: 1) contingent faculty, defined as part- and full-time faculty appointed off the tenure track and whose sole responsibility is to teach, and 2) specialized teaching focused faculty who are often hired with greater job security and are required to engage in discipline-based education research.

2.2 Contingent Faculty

One of the most prominent features of the American higher education in the past three decades is an increasing reliance on non-tenure track contingent faculty. Typically hired on short-term contracts with an institution and many on part-time basis, these faculty now account for almost three-quarters of instructional faculty at American higher education institutions (Gyurko et al., 2016). The primary responsibility of contingent faculty is to teach; each contingent faculty member typically takes on 4-5 teaching assignments per term, and contingent faculty are often assigned introductory courses with the largest enrollment size. Thus, contingent faculty relieve tenure-track faculty from some of their teaching duties, theoretically contributing to improved research-teaching balance at institutions.

The heavy reliance on contingent college instructors have resulted in national debate regarding the impacts of these faculty on student academic performance. On the one hand, many of the contingent faculty are previously or concurrently employed in the private sector in their respective fields and are therefore able to make content applicable to the 'real world.' Through their network with the private sector, contingent faculty may also help students secure internships and jobs, thus potentially enabling students to make smoother transitions to the labor market. Qualitative research also noted that when provided with professional development opportunities, contingent faculty are likely to apply evidence-based teaching practices in their own teaching, indicating that these faculty are willing to make efforts to improve their instructional effectiveness (Gyurko et al., 2016). On the other hand, however, there is widespread concern regarding the working conditions of contingent faculty. Researchers noted that contingent faculty are generally assigned heavy teaching loads that are often split between campuses; most of the departments do not actively engage contingent faculty in departmental decision making process and curriculum development; contingent faculty often do not have office space on campus and have limited knowledge regarding the student service resources on campus, therefore substantially limiting their capacity in forming a strong and interactive academic relationship with students (Kezar & Maxey, 2014; Schuster, 2003; Umbach, 2007).

A handful of studies have shed light on this issue by explicitly comparing instructional practices and interactions with students between different types of faculty. Umbach (2007), for example, used data from the Faculty Survey of Student Engagement to assess whether contingent faculty utilize sound pedagogical practices in undergraduate education when compared with tenure-track faculty practices. The specific practices included in the analysis were faculty use of active and collaborative learning techniques, faculty effort to convey high expectations in terms of student assignments, faculty time spent preparing for class, and faculty interactions with students. The author found that part-time contingent faculty interacted with students less often, used active and collaborative instructional strategies less frequently, spent less time preparing for class, and had lower academic expectations for their students than did their tenured and tenure-eligible counterparts. Full-time contingent faculty, however, looked more like their tenure track counterparts; for example, they used active and collaborative instructional techniques similarly. They also spent more time preparing for class than tenure track faculty did.

Using the 2004 National Study of Postsecondary Faculty, Baldwin and Wawrzynski (2011) examined faculty type and pedagogical approaches to learning. The specific approaches observed in this study were the use of learning-centered approaches (e.g., students' evaluations of each other's work, essay midterm/final exams, term/research papers, and students submitting multiple drafts of written work) and subject-centered teaching strategies (e.g., multiple choice midterm/final exams and short answer midterm/final exams). Their findings echo those of Umback (2007): full-time contingent faculty approached their teaching in a similar way as their tenure track counterparts did.

Specifically, they employed learning-centered strategies in the majority of their classes with greater frequency than part-time contingent faculty did. Compared with part-time contingent faculty, both tenure track and full-time contingent faculty were more likely to use technology, such as email platforms and websites, allowing better communication and interaction with students.

While these studies make an important first step toward understanding how the instructional approaches may differ by different types of faculty, none of them is able to directly link types of instructor with student performance outcomes. More recently, with the increasing access to college administrative data, a growing number of studies have used rigorous quasi-experimental designs to compare student outcomes by contingent versus tenure track faculty (e.g., Bettinger & Long, 2010; Carrell & West, 2010; Ehrenberg & Zhang, 2005; Figlio et al., 2015; Jaeger & Hinz, 2008; Ronco & Cahill, 2006; Ran & Xu, in press). Yet, the results are mixed and far from conclusive. One possible explanation for the inconsistent findings across studies might be the wide variations in contingent faculty's employment contracts and relationship with the institution. For example, Figlio et al. (2015) explored the impact of contingent teaching faculty on students' course enrollment and success relative to tenure-track/tenured faculty, drawing upon data from eight recent cohorts of first-year students at Northwestern University. The authors noted that the majority of the contingent faculty at Northwestern had a long-term relationship with the university and many were hired on a full-time basis. Based on a two-way fixed effects model that controlled for student fixed effects and next class fixed effects, the authors identified a sizable positive impact of contingent faculty on students' subsequent performance.

In contrast, using data that include both public two-year community colleges and four-year institutions in an anonymous state, Ran and Xu (in press) found that non-tenure-track instructors are positively associated with students' contemporaneous course performance but are negatively associated with subsequent course enrollment and performance within the same field of study, where the negative impacts on students' subsequent outcomes are particularly pronounced among contingent faculty hired through short-term contracts. Similarly, using data from the Virginia state community college system, Xu (2019) defined contingent faculty as part-time adjunct instructors and identified a sizable negative impact of these faculty on students' subsequent enrollment and performance. These differential results imply that the impacts of contingent faculty relative to traditional tenure track faculty may vary depending on the characteristics of the faculty and their contract and specific work conditions at an institution.

2.3 Teaching-Focused Faculty (TFF)

Compared with contingent faculty who are often hired through temporary positions, teaching-focused faculty (TFF) is a more specialized category of faculty whose primary responsibility is teaching rather than research. Similar to contingent faculty, TFF are responsible for large teaching loads, but their work conditions are better, as they have the potential to gain job security and are provided additional benefits that make them a fundamental part of the university and their respective departments. For example, at many institutions, TFF attend departmental meetings, are able to vote in regard to departmental decisions, and are given office space. At a number of institutions, TFF also engage in DBER scholarship and often serve as educational leaders within their departments. TFF's responsibilities and their roles on campus thus make them well-positioned to improve undergraduate education.

Despite the high hopes surrounding TFF to improve teaching quality, however, no study has linked TFF to student outcomes relative to either research or contingent faculty. The majority of studies about TFF have been conducted in the field of STEM education, where this faculty type is known as Science Faculty with Education Specialties (SFES). These studies investigated the institutional motivation for hiring SFES, variation among SFES in terms of observable background characteristics (e.g., formal pedagogical training in science education), and their primary responsibilities and activities at an institution (Bush, Stevens, Tanner, & Williams, 2017; Bush, Pelaez, Rudd, Stevens, Tanner, & Williams, 2015; 2013).

Overall, results from these studies provide valuable information regarding how TFF are conceptualized. First, although their main role is teaching science, not all SFES have formal training in science education (Bush et al., 2013). Second, there may be misalignments between the institution's original motivation for hiring SFES and how these faculty actually function. For example, at Ph.D.-granting institutions, SFES commonly report spending the majority of their time on relieving the research faculty from teaching and service, while having limited time to participate in education research (Bush et al., 2015; 2017). These results call into question SFESs' ability to have a strong leadership role as teaching professors and the level of support to these faculty to improve undergraduate education at the institutions employing them.

On the other hand, studies outside of SFES research suggest that TFF seem likely to have greater potential to provide effective instruction than both tenured/tenure-track faculty and contingent faculty. As an example, Rawn and Fox (2018) created an extensive TFF survey that was used at research-intensive Canadian universities where the number of full-time TFF had increased by roughly 60% from 2006 to 2010. To put this statistic in perspective, in the research-focused track, tenure and tenure-eligible faculty have been employed at a rate that has increased by only 10% over the same time period. Rawn and Fox (2018) found that the majority of TFF engaged in a wide variety of work, including teaching, service, curriculum leadership, and often DBER. Among these activities, teaching undergraduate courses and department service were the two most commonly reported activities. In terms of scholarship, more respondents identified conducting research on pedagogical topics than on disciplinary topics. They also reported engaging in curriculum leadership often. These two findings distinguish TFF from both contingent faculty and traditional research faculty; they also support the idea that TFF have the potential to teach well and elevate the quality of teaching within their departments.

Harlow and Sato (2017) has undertaken an in-depth examination of TFF in the U.S. setting. Specifically, they surveyed tenure-track teaching faculty members in the UC system, the context of the present study, in an effort to generate a descriptive picture of their background, role, and overall sentiments about the profession. The study found that tenure-track teaching faculty, on average, were satisfied, felt valued, and experienced a sense of belonging in regard to their departments. Further, similar to the sample of TFF in the Rawn and Fox (2018) study, roughly 60% of TFF in the UC system participated in DBER.

Finally, based on a unique classroom observation dataset including the instructional practices used in 259 large introductory STEM courses at a single UC campus, Vu (2017) compared the teaching practices by different types of faculty. The author found that teaching faculty used more “promising practices” (e.g., formative assessment, active learning, and collaborative group activities) than both research faculty and lecturers, therefore providing some suggestive evidence about the possible benefits of TFF for student learning.

Higher education administrators must balance many goals; for example, research universities aspire to excellence in regard to *both* research and teaching. Employing specialized faculty such as TFF is one way to reach these goals given teaching faculty’s leadership role in teaching. However, the relevant literature outlined above in regard to both contingent faculty and TFF suggests that the role of teaching faculty and their instructional approaches may vary by their working conditions, resources devoted to instruction, prior pedagogical experiences, and faculty expectations. As the UC system and other higher education institutions are contemplating hiring more teaching-oriented faculty, results from the current study will inform these discussions by assessing the relative effectiveness of different types of faculty in terms of student academic outcomes.

3. Study Background

3.1 Data and Institutional Context

Our analysis is based on detailed student and administrative data from one of the University of California (UC) campuses. The UC system consists of ten campuses with a total enrollment of approximately 200,000 undergraduate students in the 2013 school year. Compared with other selective four-year colleges nationwide, the anonymous UC campus (referred to as "AUCC" here after) in our study has a diverse student population, where a large proportion of the student body belongs to an under-represented minority (URM) group.

All the data used in our study was collected from the Office of Institutional Research (OIR). The data includes six cohorts of first-time college students entering the AUCC between fall 2008 and fall 2013. The student-level data includes demographic characteristics, such as gender, race, and age, and academic preparation variables such as SAT score (see Table 3.1 for a full list). The data also includes course enrollment and transcript data, therefore providing information about the type of courses each student takes during a particular quarter as well as about the associated course instructor types and course grades.

Table 3.1 displays descriptive statistics for college students entering AUCC in fall 2008, compared with characteristics of students entering any public four-year college nationwide in the same year based on data from the Integrated Postsecondary Education Data System (IPEDS). It seems that AUCC includes a much larger proportion of Asian students: Asian students comprise 55 percent of the 2008 enrolling cohort compared with 7 percent nationally. In addition, AUCC is a relatively more selective campus with SAT-math scores over 100 points above the national average and SAT-verbal scores roughly 70 points above the national average.

3.2 Faculty Composition at AUCC

During the period of this study, AUCC divided faculty into 50 distinct faculty titles. Of those 50 title codes, half were positions with limited or no teaching responsibilities, such as visiting scholars, research specialists, academic administrative officers, and are therefore excluded from our analyses. Since the motivation of hiring teaching faculty and graduate student instructors are distinct from each other, we also exclude courses taught by graduate students.[19] The remaining faculty categories relevant to our study fall under three broad faculty categories: tenure-track research faculty, contingent lecturers, and tenure-track teaching faculty.[20] We briefly describe the job responsibilities of each category below.

Tenure-track research faculty in the UC system belong to the Ladder-series. Faculty in this group include assistant professors, associate professors, and full professors. All of them are part of the academic senate, where faculty carry out shared governance responsibilities established by the UC governing body ("The Board of Regents"). For example, they are involved in academic policy, admissions and granting of degrees, authorization and supervision of new courses and curricula, and are also responsible for the administration on faculty appointments, promotions, and budgets.

Contingent faculty in the UC system comprise of lecturers and adjuncts. These are non-academic senate positions that are offered both part-time and full-time. Lecturers are hired to provide instructional services. They may teach up to 9 courses over a 3-quarter

[19] The total course enrollments with graduate students are 14% of the student-course sample.
[20] The complete list of faculty titles and the corresponding broad faculty categories are presented in Appendix Table 3.1.

year and are often assigned to teaching introductory courses with large enrollment size. Compared with the lecturer positions, the responsibility of adjunct faculty can be more flexible, where they can be hired to either teach courses or conduct research. The other key distinction is that adjunct faculty can be considered for advancement through the professorial series which includes: adjunct instructor, adjunct assistant professor, adjunct associate professor, and adjunct professor. In our dataset, only a small proportion of courses were taught by adjunct instructors (less than one percent of total course enrollments in our first-term course analytical sample). Given this, and considering that they are a heterogeneous group of instructors, we have removed adjunct instructors from all analyses. In a separate robustness check, we combine lecturers and adjuncts into one category and the results remain almost identical.

Lastly, tenure-track teaching faculty in the UC system belong to the Lecturer with Security of Employment (LSOE)-series. These positions are designed to meet the long-term instructional needs of a university, where teaching faculty's primary responsibilities are teaching, and their secondary responsibility is to engage in DBER. In some departments, teaching faculty are expected to take leadership roles in the areas of teaching, instructional development, curriculum design, course structure, teaching methods, and new technologies. They are also required to coordinate a spectrum of professional development teaching activities. Similar to contingent lecturers, teaching faculty are responsible for large teaching loads, typically teaching two large undergraduate courses per quarter, totaling five to six courses per year. Yet, as part of the tenure system, teaching faculty are also Academic Senate members and have the same rights and privileges in the departments and on the campus as Senate research faculty. For example, teaching faculty have full voting rights in terms of departmental decisions, are eligible for appointment to Senate committees, and are evaluated for re-appointment, merit salary increases, and promotion to tenure in the same manner as research faculty in the UC system. However, the evaluation criteria differ for teaching faculty and research faculty, where teaching faculty are largely evaluated based on their teaching excellence and professional achievements, while research productivity and excellence is paramount in terms of promotion decisions for research faculty.

It should be noted that while the term LSOE has also been used outside of the UC system to refer to non-tenure-track lecturers with long-term employment with an institution, the LSOE-series in the UC system refers specifically to teaching faculty who are members of the Academic Senate and have achieved (or have the potential to achieve) tenure for teaching. As such, there is a roughly parallel track between teaching faculty and research faculty: Lecturer with Potential Security of Employment (Lecturer PSOE) are at the same level as assistant professors; Lecturer with Security of Employment (Lecturer SOE) are at the same level as associate professors; senior Lecturer with Security of Employment (Senior Lecturer SOE) are at the same level as full professors.[21] Promotion through the teaching faculty-series is fairly similar to the traditional research faculty-series where an assistant teaching professor becomes an associate teaching professor and an associate teaching professor becomes a full teaching professor.

---

[21] The term "Senior", whether appointed to a PSOE position or SOE position, is assigned to individuals who provide services of exceptional value to the University and whose excellent teaching and professional accomplishments have made him or her a recognized leader in his or her professional field and/or in education.

Table 3.2 presents the average characteristics of the three categories of faculty among those who taught at least one course between fall 2008 and summer 2016, which includes 80 teaching faculty, 1,329 research faculty, and 1,090 lecturers (columns 1-3). Compared with research faculty, teaching faculty are substantially more likely to be female (50% vs. 35%), white (78% vs. 68%), and less likely to be Asian (9% vs. 20%). Other than these demographic differences, however, teaching faculty and research faculty are fairly comparable in their age, full-time employment status, and total years working at the university. Yet, the average monthly pay rate is noticeably lower for teaching faculty compared with research faculty ($7,021 vs. $10,391) despite much heavier teaching load among teaching faculty than that of research faculty (2.3 courses per term vs. 1.5 courses per term or 8.2 course credits per term vs. 5.5 course credits).[22]

Compared with both types of tenure-track faculty, lecturers at AUCC tend to be younger and have fewer years as an employee at the university. Four-fifths of the lecturers are also employed on a part-time basis with a monthly pay rate half the rate for tenure-track research faculty. Although the majority of lecturers are part-timers, their teaching load is roughly equivalent to research faculty (1.5 courses per term or 5.5 course credits per term).[23] Columns 4-6 further limit the sample to faculty who taught at least one course in our analytical sample, which only includes the first course a student took in a field of study. The patterns of the between-faculty distinction resemble the patterns described for the full-course sample.

---

[22] It should be noted that given that the monthly pay rate includes faculty responsibilities beyond teaching, it is difficult to get an accurate measure of the gap in average per-course-credit income.
[23] A detailed explanation regarding benefits eligibility can be found at the AUCC website: https://ucnet.universityofcalifornia.edu/tools-and-services/administrators/docs/upay726-beli-form.pdf

To further shed light on possible variations in reliance on different types of faculty across departments, Figure 1 shows the proportions of total student course enrollments with each type of faculty by fields of study using the analytical sample.[24] In general, research faculty serve as the primary teaching force in the majority of fields. In some fields, such as Humanities, almost all the course enrollments are with research faculty. Tenure-track teaching faculty are responsible for a relatively large number of course enrollments in four of the twenty-two distinct fields of study: Biology, Chemistry, Computer Sciences, and Social Ecology. In two of these subfields, Social Ecology and Chemistry, teaching faculty are responsible for 55% and 31% of course enrollments, respectively. In contrast, lecturers are less likely to be housed in STEM fields, with the exception of being responsible for roughly 60% of course enrollments in mathematics. Lecturers also teach a relatively large proportion of students in English, psychology, and other social sciences.

3.3 Sample and Summary Statistics

Because the aim of this study is to evaluate the impact of instructor type during a student's *initial* exposure to a field of study on the student's concurrent and subsequent academic outcomes, we limit the analysis to the first course a student takes in each field of study (referred to as "introductory course" hereafter). We further limit the sample to courses taken during a student's first quarter at AUCC to minimize the possibility of student systematic sorting by type of faculty, as this is the time when students are least likely to have existing knowledge about different types of faculty. The final analytical

[24] The specific number of enrollments with each type of faculty across fields of study is presented in Appendix Table 3.2.

sample includes 80,980 first-term introductory course enrollments for 36,965 students. These enrollments are from 1,075 courses taught by 1,076 distinct faculty members.

In estimating the impact of faculty in first-term introductory courses, we consider both current and subsequent academic outcomes in the same field of study. Specifically, we begin our inquiry by examining whether instructor type—teaching faculty vs. research faculty vs. lecturer—impacts contemporaneous student course outcomes, specifically examining a) course grade on a zero to four grading scale for all students and among students who successfully passed the course.[25] Yet, results of current course performance are difficult to interpret without additional information. For example, a significant positive impact from taking introductory courses with teaching faculty might be explained by either more effective instruction or simply different grading patterns. Therefore, we further use measures beyond the first course students took in a field of study to provide a more comprehensive understanding of the impact of different types of faculty on student outcomes. Specifically, we use subsequent course enrollment—whether students enroll in a second course in the same field that is different from the introductory course after initial exposure—to measure student interests in a subject. Prior research has also noted that subsequent performance is an important measure of student learning, as it indicates whether instructional quality in introductory courses has a lasting impact on learning and engagement in the same field of study (Carrell & West, 2010). We therefore also include students' performance in the second course within the same field of study as an outcome measure. Finally, we also examine students' *longer-term* interest in a field and include as

[25] We did not include course persistence as an outcome measure in this study because almost no student in our sample withdrew from a course, resulting in an overall course persistence rate of 99.7%.

outcome measures the total number of credits students earn subsequently in the same field of study and whether students declare a major in that field.

Table 3.3 provides summary statistics for the outcome measures: Panel A presents contemporaneous course outcomes, Panel B presents subsequent enrollment outcomes, and Panel C presents subsequent academic outcomes. As shown in Panel A, average course grades are fairly similar among the three types of faculty. In terms of subsequent enrollment and performance outcomes, the overall probability that a student takes any additional courses within the same field of study is 84%. Students who take their introductory course with teaching faculty or with research faculty seem to be equally likely to enroll in a second course in the same field of study, and both are higher compared with students who take an introductory course with a lecturer. In terms of subsequent performance, students who take an introductory course with teaching faculty earn the lowest grades, on average, in subsequent courses. However, as mentioned previously, these descriptive patterns could reflect student-level and course-level selection. Lastly, students who take their introductory course with teaching or research faculty earn roughly 25 to 26 credits in the same field of study, both of which are substantially higher compared with students who take introductory courses with lectures (18 credits). Similarly, students also seem to be least likely to declare a major in the same field of study as their introductory course when it is taken with lecturers (17%), compared with 30% for research faculty and 20% for teaching faculty.

4. Estimation Strategy

4.1 Empirical Model for Contemporaneous Course Outcomes and Subsequent Enrollment

Our primary empirical specification for current course outcomes and subsequent enrollment outcomes relates student *i*'s outcomes (*Y*) in section *s* of course *c* in field *k* in term *t* to the type of instructor that the student had during his/her initial exposure to a field ofstudy:[26]:

$$Y_{icskt} = \alpha + \beta \text{ Instructor}_{icskt} + \rho_{ck} + \pi_t + X_{cskt} + \gamma_i + \mu_{icskt} \quad (1)$$

The key explanatory variable is the type of instructor with whom a student took the introductory course. We use lecturer as the reference group and therefore the vector ($\text{Instructor}_{icskt}$) includes two dichotomous variables: research faculty and teaching faculty. ρ_{ck} represents course fixed effects, which captures both observable and unobservable course-level characteristics (e.g., field of study, pre-requisites, course requirements, course topics, and level of difficulty). The course fixed effects allows for comparisons among different sections of the same course taught by different types of instructors (e.g., Introductory Microeconomics). π_t represents term fixed effects that takes into account overall variations in course grades and faculty composition over time at AUCC. In addition to the term fixed effects and course fixed effects, the model also controls for student-term-level information that varies over time (i.e., total credits taken in the term) and characteristics that are specific to a particular section (i.e., whether the course section is

[26] We use the term "section" or "class" to refer to a particular offering of a course with a specific instructor at a specific time, such as "MATH 101 Calculus – section 01 in the fall term of 2009." Hence, a "section" or "class" is uniquely defined by course title, section, and term.

online or face-to-face, the section enrollment size, the average SAT score of peers enrolled in the section to capture peer effects, and day and time of the class), as indicated by X_{cskt}.

The remaining source of selection is due to students' individual preferences for different types of instructors within courses. For example, more academically motivated students might prefer research faculty for potential research opportunities. It is worth noting that the extent of student sorting is already limited given that we only focus on the courses taken during a student's initial quarter at AUCC. But we can also investigate this question directly by relating the type of instructor teaching a section to a wide range of student-level characteristics controlling for course and term fixed effects, and observable section-level characteristics.

The results presented in Table 3.4 suggests that while there is no consistent relationship between types of instructors and indicators of student previous academic performance, low-income students seem to sort into courses differently based on faculty teaching the course. The point estimates show that they are more likely to take their introductory courses with research faculty compared with their high-income counterparts, and less likely to take their introductory courses with teaching faculty. To address possible selection bias due to student sorting by types of instructors, we further include student fixed effects (γ_i) into the model, which controls for both observable and unobservable student-level characteristics that are constant for an individual, such as students with certain characteristics having preferences for a particular type of instructor.[27]

[27] It should be noted that due to the course fixed effects, only courses that have been taught by more than one type of faculty would contribute to the estimation of instructor impact on student outcomes. In a similar vein, as a result of student individual fixed effects, a student who only takes her introductory courses with a certain type of faculty (e.g., research faculty) would not contribute to the estimation of the instructor effects either. Yet, these courses and students still contribute to the precision of the estimates for other variables in the model, such as the time fixed effects. Therefore, it is commonly recommended by researchers to keep these

Table 3.4 also includes section-level characteristics, such as class size, whether the course is upper or lower division, whether the section is delivered through the online or face-to-face format, and whether the section is offered in the morning or afternoon/evening. Two associations stand out. First, even though 92% of the courses in the analytical sample are lower-division courses, research faculty are less likely to teach them, as compared with lecturers. Second, teaching faculty are less likely to teach courses that are offered online. Given the association between section-level characteristics and faculty type, we include them in all models as indicated by X_{cskt}. Equation (1) draws on two sources of variation. The first includes student-level variations, whereby a student takes introductory courses with different types of instructors in different fields of study. For example, a student may take an introductory physics course with a teaching faculty but an introductory math course with research faculty. The vast majority (95%) of the students in our first-term freshman sample took their introductory courses with a mixture of different types of instructors. In addition to within-student variation, we also draw on between-section variations in the type of instructors within a course. This could be either due to within-term variations, where multiple sections are offered during a particular term and are also taught by different types of faculty in that term (14% of the course enrollments are from courses with such variation) or to between-term variation, where different types of instructors teach the same course during different terms (70% of the enrollments are from courses with such over-time variation). The fact that we have sufficient variation both within individual students and within courses in types of faculty thus supports the use of

observations in the sample (Wooldridge, 2002). Appendix Table 3.3 shows the characteristics of students who contribute to the fixed effects estimators, which are not substantially different from the characteristics of students in the full sample.

the two-way fixed effects model. We cluster standard errors at both the student and at the subject level to take into account correlations within classes taken by the same student, as well as correlations between classes within the same field of study.

4.2 Empirical Model for Subsequent Course Outcomes

Equation (1) enables us to estimate the impact of taking one's introductory course with different types of faculty on students' performance in the introductory course as well as their subsequent enrollment in the same field of study. However, when examining the impact of introductory course instructors on the performance in the subsequent course, we are concerned that initial experiences may also influence next class choice. For example, if a student finds the introductory course too difficult, she might opt into a less challenging course for her subsequent enrollment in the same field. In a similar vein, if a student took her introductory course with a lecturer and also had unsatisfactory experience in that course, she may intentionally avoid classes taught by that instructor or even all classes taught by lecturers in the same department.

To address the additional variation in subsequent class choice, we extend equation (1) by further controlling for next-class fixed effects for subsequent course performance measures:

$$Y_{icskt+1} = \alpha + \beta\ \text{Instructor}_{icskt} + S_{cskt+1} + \gamma_i + \rho_{ck} + \pi_t + X_{cskt} + \mu_{icskt+1} \qquad (2)$$

where student i's outcomes in the next class section s in course c in field k in time $t+1$ are related to the student's experience with the type of instructor that the student had in his/her introductory course ($\text{Instructor}_{icskt}$). Since the next-class fixed effect S_{cskt+1} is a combination of course, time, and specific section, this model specification compares student performance in exactly the same next course-section, therefore controlling for the possibility that initial experience in a field may influence a student's next course choice and preference for different types of instructors in the next course in the field. As a result, the variation for equation (2) comes from students taking exactly the same course-section as their second class in a field (S_{cskt+1}), where some students would have had teaching faculty in their prior introductory course in that field while others would have taken the same introductory course with research faculty or lecturers.

It is important to note that the introductory course fixed effects are still in the model. Controlling for characteristics at the introductory course level is necessary even though the outcome is a measure of student performance in the subsequent course. For example, a student might take an introductory course, such as microeconomics taught by a teaching faculty, which prepares him well for a subsequent course, macroeconomics. A different student, however, may take an introductory course—the history of economics, for example—with a research faculty member, thereby becoming ill-prepared for macroeconomics. If the student taking microeconomics does better (or worse) in macroeconomics, this impact can be the result of the student taking a class that better prepared him for the subsequent course. Therefore, it is necessary to control for introductory course differences for examining either current or subsequent outcomes.

5. Results

5.1 Contemporaneous Course Outcomes

Table 3.5 presents the contemporaneous course results based on equation (1) that controls for both student fixed effects and introductory course fixed effects, and the outcome measure is the numeric measure of course grade on a 0 to 4 grading scale, where

A corresponds to 4 and F corresponds to 0. Column 1 includes students who withdrew from courses, where a withdrawal corresponds to a 0 and column 2 excludes these students from the analysis.

The results indicate that there are no significant differences among students who take their introductory courses with research faculty or teaching faculty, as compared with lecturers. Given that average effects may mask potential differences in grade distribution, we also explore the impact of different types of instructors on the full distribution of the letter grades in introductory courses. Indeed, the results presented in Table 3.6 indicate that while tenure-track teaching faculty are equally like to grant As, Bs, and Cs compared with contingent lecturers, students on the margin of failing the course are significantly more likely to receive an F by roughly 3 percentage points when taking the course with tenure-track teaching faculty.[28]

5.2 Subsequent Outcomes

Table 3.7 presents results for subsequent outcomes. Column 1 measures whether a student takes on another course in the same field of study after her initial introductory course. Column 2 uses the average course grade of the next course a student takes (excluding the student's own course grade) to capture the possible influence of initial experiences with the introductory course on the difficulty of a student's subsequent course choice in the same field. Additionally, column 3 presents the results for subsequent course grade based on equation (2) that controls for student fixed effects, introductory course fixed effects, and next-class fixed effects, therefore comparing students who take the same

[28] The coefficient remains negative and significant when we change the reference group from lecturers to research faculty, indicating that students are more likely to receive an F when they take an introductory course with teaching faculty compared with either research faculty or lecturers.

introductory course, as well as the same course-section for the second class enrollment in that field. Lastly, columns 4 & 5 show the results for our two longer-term outcome measures of student interest: total credits earned in the same field of study and whether students declare a major in the same field of study as their introductory course.

Similar to the results for contemporaneous course outcomes, we do not identify significant differences between research faculty and lecturers in terms of student subsequent course enrollment and performance. However, students taking their introductory courses with teaching faculty earn subsequent course grades that are 0.05 grade units higher than students taking their introductory course with lecturers. In terms of longer-term measures of student interest, we also find a small and marginally significant impact of teaching faculty on total subsequent credits earned in a field of study, suggesting that teaching faculty may have the potential to impact students' longer-term outcomes in the field. We investigate this possibility further by estimating the impact of faculty type on three different outcome measures: (i) the probability of earning at least one more credit after the initial introductory course, (ii) the probability of earning eight or more credits subsequently after the initial introductory course (equal to two more four-credit courses), and (iii) the probability of earning twelve or more credits subsequently (equal to three or more four-credit courses). The results are presented in Appendix Table 3.4. Students taking their introductory course with teaching faculty, as compared with lecturers, are equally likely to earn at least 1 credit in the same field of study following the introductory course. This echoes the null finding we reported above regarding subsequent enrollment (Table 3.7). Yet, students taking their introductory course with teaching faculty are 4 percentage points more likely to earn at least eight credits and at least twelve credits subsequently in the same field of study. These findings seem to suggest that students who take their introductory course with teaching faculty are eventually more likely to take a third and even fourth course in the same field.

6. Discussion and Conclusion

Research universities aim for excellence with regard to both research and teaching, striving to create a diversified workforce with varying job responsibilities to fulfill these goals. To meet the rising demands of undergraduate student enrollment and increased teaching obligations, colleges and universities have increasingly relied on instructors whose primary responsibility is to teach. Hiring teaching faculty protects the scholarship time of those who are traditionally responsible for both teaching and research, potentially improving overall production within a department or school. However, the dual institutional goals of achieving excellence in research and teaching cannot be fully met unless teaching faculty—those that are contingent or on the tenure-track—are as effective as traditional research faculty. Examining student outcomes as a function of faculty type therefore provides an empirical foundation for helping policymakers and higher education administrators better understand the consequences of their staffing decisions.

This paper examines two distinct teaching faculty models: i) contingent faculty, defined as either part- or full-time faculty members who are appointed off the tenure track with the primary role of teaching, and ii) UC-specific tenure-track teaching faculty, defined as faculty whose primary role is to teach and whose secondary role is to engage in scholarship activities, such as DBER. Using data from six cohorts of students enrolled at AUCC, we compare both tenure-track research faculty and teaching faculty with contingent faculty in terms of their impacts on current and subsequent student academic outcomes. Our results indicate that the three types of faculty are fairly comparable for most of the student outcome measures. Yet, students who took their introductory course with tenure-track teaching faculty seem to earn slightly better subsequent course grades and accumulate more subsequent credits than otherwise similar students who took their introductory course with lecturers or research faculty.

These results provide several implications for institutional faculty hiring policies. First, as researchers and college administrators have been engaged in discussions about whether institutions should grant tenure appointment to teaching faculty, our results suggest that a tenure-track teaching faculty model may indeed serve as an effective way to professionalize and stabilize the teaching workforce in higher education. Compared with their tenure-track research colleagues and lecturers, tenure-track teaching faculty in our sample assume a substantially larger teaching load and produce comparable and sometimes even better student learning outcomes than either tenure-track research faculty or contingent lecturers. These results therefore nominate tenure-track teaching faculty as a promising alternative to research faculty at least in teaching introductory coursework. This nomination is in line with a handful of qualitative studies that further advocate for the broad benefits of tenure-track teaching faculty to an institution. For example, Bush and his colleagues (2015) described the potential of teaching faculty to serve as pedagogical leaders and agents of change within their respective departments. Harlow, Buswell, Lo, & Sato (2019) also noted that tenure-track teaching faculty are more economical than other types of faculty because they not only take on large teaching loads but also engage in service and are often the ones to lead student service programs and take on other administrative duties.

Second, our results are also associated with a small but growing body of literature concerning the impact of contingent faculty on student learning outcomes. At colleges and universities across the country, employing contingent faculty has become a way to fulfill increasing student enrollments and teaching obligations. Today, more than two-thirds of the professoriate in non-profit postsecondary education is comprised of contingent faculty (Kezar & Maxey, 2013). Considering that the productivity of contingent faculty may depend largely on their working conditions at an institution, it is important to assess the effects of contingent faculty under different institutional contexts. The field thus far has accumulated evidence on this topic at private elite four-year institutions (Figlio et al., 2015), public elite four-year institutions in Canada (Hoffmann and Oreopoulos, 2009), less-selective public four-year institutions (Bettinger & Long, 2010; Ran & Xu, in press), and open-access two-year community colleges (Ran & Xu, in press; Xu, 2019).[29] Our study contributes to this literature by estimating the effects of contingent faculty at a new setting—an elite U.S. public four-year institution. Our results in regard to the impact of contingent faculty on student outcomes are similar to those of Hoffmann & Oreopoulos (2009) that also did not find significant differences in student outcomes between contingent lecturers and tenure-track research faculty.

Our results also point to several directions for future research. First, without detailed information on instructional practices, grading criteria, and instructor-student interactions, the current study is not able to identify the specific mechanisms through

[29] See Appendix table 3.5 for the key findings from the current study and from the existing literature that directly compared student outcomes between contingent faculty and tenure-track faculty.

which different types of faculty influence student outcomes. One possible explanation for the positive relationship between tenure-track teaching faculty and student subsequent outcomes is that tenure-track teaching faculty use instructional strategies that stimulate interest in a field or, at minimum, provide students with a relatively enjoyable classroom experience. A related possibility is that teaching faculty are more actively involved in student service programs or provide more interactive opportunities to students after class, which may lead to closer instructor-student relationships that prove to be critical to students' motivation and interests (Crockett, 2015; Easton & Guskey, 1983; Kezar & Maxey, 2014; Pascarella & Terenzini, 1995, 2005; Raisman, 2013; Umbach & Wawrzynski, 2005). Future research may wish to collect detailed classroom observation data, student interview data, and instructor interview data to shed light on these possible operating channels.

Additionally, considering that AUCC resembles other UC campuses and public research institutions in terms of faculty obligation and student population, the educational benefits of tenure-track teaching faculty may well apply to other universities. Future research may wish to compare the relative effect of tenure-track teaching faculty in terms of student learning outcomes on other UC campuses, as this can likely provide insight regarding the generalizability of our findings. Future research is also particularly important in this regard given the small pool of tenure-track teaching faculty, relative to other faculty types, in our sample. Lastly, it is also important to note that since our analyses focus only on the first course a student takes in a field, whether these results are likely to extend to more advanced coursework still requires future research.

References

Astin, A. (1993). *What matters in college: Four critical years revisited.* San Francisco: Jossey-Bass.

Baldwin, R.G. & Wawrzynski, M.R. (2011). Contingent faculty as teachers: What we know; what we need to know. *American Behavioral Scientist, 55*(11). 1485-1509.

Benjamin, E. (2002). How over-reliance on contingent appointments diminishes faculty involvement in student learning. *Peer Review, 5*(1).

Benjamin, E. (2003). Reappraisal and implications for policy and research. *New Directions for Higher Education, 123,* 79-113.

Berube, M. & Ruth, J. (2015). *The humanities, higher education, and academic Freedom: Three necessary arguments.* New York, New York: Palgrave Macmillan.

Bettinger, E. P., & Long, B. T. (2010). Does cheaper mean better? The impact of using adjunct instructors on student outcomes. *Review of Economics and Statistics, 92*(3), 598–613.

Braxton, J. M., Bray, N. J., & Berger, J. B. (2000). Faculty teaching skills and their influence on the college student departure process. *Journal of College Student Development*, *41*, 215-227.

Brew, A. (1999). Research and teaching: Changing relationships in a changing context. *Studies in Higher Education*, 24(3), 291-301.

Brew, A. (2013). Understanding the scope of undergraduate research: A framework for curricular and pedagogical decision-making. *Higher Education, 66,* 603-618.

Brew, A. & Jewell, E. (2012). Enhancing quality learning through experiences of research-based learning: Implications for academic development. *International Journal for Academic Development, 17,* 47-58. DOI:

Bush, S. D., Pelaez, N. J., Rudd, J. A., Stevens, M. T., Tanner, K. D., & Williams, K. S. (2013). Widespread distribution and unexpected variation among science faculty with education specialties (SFES) across the United States. *Proceedings of the National Academies of Sciences of the United States of America*, 110(18), 7170-7175.

Bush, S. D., Pelaez, N. J., Rudd, J. A., Stevens, M. T., Tanner, K. D., & Williams, K. S. (2015). Misalignments: Challenges in cultivating science faculty with education specialties in your department. *BioScience, 65*(1), 81-89.

Bush, S. D., Stevens, M. T., Tanner, K. D., & Williams, K. S. (2017). Origins of science faculty with education specialties: Hiring motivations and prior connections explain institutional differences in the SFES phenomenon. *BioScience, 67*, 452–463.

Carrell, S. E., & West, J. E. (2010). Does professor quality matter? Evidence from random assignment of students to professors. *Journal of Political Economy, 118*(3), 409–32.

Coate, K., Barnett, R., & Williams, G. (2001). Relationships Between Teaching and Research in Higher Education in England. *Higher Education Quarterly*, 55(2), 158-174.

Condon, W., Iverson, E. R., Manduca, C. A., Rutz, C., & Willett, G. (2016). *Faculty development and student learning: Assessing the connections*. Bloomington: Indiana University Press.

Crockett, D. S. (2015). *The ten most effective retention strategies for community/technical* *colleges* [PowerPoint slides]. Retrieved from http://docplayer.net/3064806-The-ten-most-effective-retention-strategies-for-community-technical-colleges-david-s-crockett.html

Easton, J.Q. & Guskey, T.R. (1983). Estimating the effects of college, department, course, and teacher on earned credit rates. *Research in Higher Education, 19,* (153-158).

Ehrenberg, R. G., & Zhang, L. (2005). Do tenured and tenure-track faculty matter? *Journal of Human Resources, 40*(3), 647–59.

Figlio, D. N., & Shapiro, M. (2017, January 26). *Are great teachers poor scholars?* Evidence Speaks Reports. Retrieved from https://www.brookings.edu/research/are-great-teachers-poor-scholars/

Figlio, D. N., Schapiro, M. O., & Soter, K. B. (2015). Are tenure track professors better teachers? *Review of Economics and Statistics, 97*(4), 715-724.

Finkelstein, B. (1984). Education and the retreat from democracy in the United States. *Teachers College Record, 20*, 275-282.

Gaff, J. G., Pruitt-Logan, A. S., Sims, L. B., & Denecke, D. D. (2003). *Preparing future faculty in the humanities and social sciences: A guide for change*. Retrieved from the Preparing Future Faculty website: http://www.preparing-faculty.org/PFFWeb.FrontMatter.PFF4Manual.pdf

Gordon, L.M. (2012). Good teaching matters, teachers matter, and teacher education matters [lecture]. Retrieved from http://files.eric.ed.gov/fulltext/ED538614.pdf.

Gyurko, J., MacCormack, P., Bless, M. M., & Jodl, J. (2016). *Why colleges and universities need to invest in quality teaching more than ever: Faculty development, evidence-based* *teaching practices, and student success*. Retrieved from the Association of College and University Educators website: http://acue.org/quality-teaching/

Hajdarpasic, A., Brew, A., & Popenici, S. (2015). The contribution of academics' engagement in research to undergraduate education. *Studies in Higher Education, 40,* 644-657.

Hattie, J., & Marsh, H. W. (1996). The relationship between research and teaching: A meta-analysis. *Review of Educational Research*, 66(4), 507-542.

Halliwell, J. (2008). *The Nexus of Teaching and Research: Evidence and Insights from the Literature.* Higher Education Quality Council of Ontario Toronto, ON.

Harlow, A., & Sato, B. (2017). *Gaining insight into the UC system's tenure-track lecturer position*. Poster session presented at SABER.

Harlow, A., Buswell, N., Lo, S., & Sato, B. *Higher education administrators on the role of teaching (faculty) at R1 institutions.* Paper presented at NARST.

Hoffman, F., & Oreopoulos, P. (2009). Professor qualities and student achievement. *The Review of Economics & Statistics, 91*(1), 83-92.

Jaeger, A. J., & Hinz, D. (2008). The effects of part-time faculty on first semester freshmen retention: A predictive model using logistic regression. *Journal of College Student Retention: Research, Theory & Practice*, *10*, 265-286.

Kezar, A. & Maxey, D. (2013). The changing academic workforce. *Trusteeship Magazine, 21*(3), 15-21.

Kezar, A. & Maxey, D. (2014). Faculty matter: So why doesn't everyone think so? *Thought & Action,* 29-44.

Kuh, G.D., & Hu, S.P. (2001). The effects of student-faculty interaction in the 1990s. *Review of Higher Education, 24*(3). 309-332.

Loes, C. N., & Pascarella, E. T. (2015). The benefits of good teaching extend beyond course achievement. *Journal of the Scholarship of Teaching and Learning, 15*(2), 1–13.

National Center for Education Statistics. (2016). Characteristics of postsecondary faculty. Retrieved from http://nces.ed.gov/programs/coe/indicator_csc.asp

Oolbekkink-Marchand, H. W., Van Driel, J. H., & Verloop, N. (2014). Perspectives on teaching and regulation of learning: A comparison of secondary and university teachers. *Teaching in Higher Education, 19*(7), 799–811.

Pascarella, E.T., & Terenzini, P. (1991). *How college affects students: Findings and insights from twenty years of research.* San Francisco: Jossey-Bass.

Pascarella, E.T., & Terenzini, P. (2005). *How college affects students: A third decade of research.* San Francisco: Jossey-Bass.

Raisman, N. (2013). The cost of college attrition at four-year colleges and universities. Virginia Beach, VA: The Educational Policy Institute.

Ran, F. & Xu, D. (in press). Does contractual form matter? The impact of different types of non-tenure track faculty on college students' academic outcomes. *Journal of Human Resources.*

Rawn, C.D. & Fox, J.A. (2018). Understanding the work and perceptions of teaching focused faculty in a changing academic landscape. *Research in Higher Education, 59*, 591-622.

Robertson, J. & Bond, C. (2001). Experiences of the relation between teaching and research: What do academics value? *Higher Education Research & Development, 20,* 5-19.

Seymour, E., & Hewitt, N. M. (1997). *Talking about leaving: Why undergraduates leave the sciences*. Boulder, CO: Westview Press.

Schuetz, P. (2002). Instructional practices of part-time and full-time faculty. In C. L. Outcalt (Ed.), Community college faculty: Characteristics, practices, and challenges. *New Directions for Community Colleges*, 118, 39–46.

Schuster, J. H. (2003). The Faculty Makeover: What Does It Mean for Students? In E. Benjamin (Ed.), *Exploring the Role of Contingent Instructional Staff in Undergraduate Learning.* San Francisco, CA: Jossey-Bass.

Teichler, U., Arimoto, A., & Cummings, W. K. (2013). *The Changing Academic Profession. Major Findings of a Comparative Survey.* Dordrecht: Springer

Tinto, V. (2004). *Student retention and graduation: Facing the truth, living with the consequences* (Occasional Paper No. 1). Washington, DC: Pell Institute for the Study of Opportunity in Higher Education.

University of California, Office of the President (UCOP) (2002). Academic personnel manual. Retrieved from: https://www.ucop.edu/academic-personnel-programs/academic-personnel-policy/.

Umbach, P. D. (2007). How effective are they? Exploring the impact of contingent faculty on undergraduate education. *Review of Higher Education, 30*(2), 91–124.

Umbach, P.D. & Wawrzynski, M.R. (2005). Faculty do matter: The role of college faculty in student learning and engagement. *Research in Higher Education, 46*(2), 153-184.

uz Zaman, M. Q. (2004). *Review of the Academic Evidence on the Relationship between Teaching and*

Research in Higher Education. Department for Education and Skills London.

Vu, V. Q. (2017). *Documenting instructional practices in large introductory STEM lecture courses* (Doctoral dissertation). Retrieved from https://escholarship.org/uc/item/1b15t5q8

Wang, J. S., Pascarella, E. T., Nelson Laird, T. F., & Ribera, A. K. (2015). How clear and organized classroom instruction and deep approaches to learning affect growth in critical thinking and need for cognition. *Studies in Higher Education, 40*(10), 1786–1807.

Xu, D. (2019). Academic performance in community colleges: The influences of part-time and full-time instructors. *American Educational Research Journal*.

Tables & Figures

Table 3.1
Student Descriptive Statistics

	National Sample	2008 Cohort	AUCC-Analytical Sample
Female	0.547	0.545	0.549
White	0.632	0.244	0.203
Black	0.108	0.022	0.027
Hispanic	0.096	0.154	0.206
Asian	0.068	0.552	0.554
Other	0.010	0.028	0.011
Student Age	---	19.094	19.290
Low-income Status	---	0.223	0.317
First-generation Status	---	0.350	0.450
HS GPA	---	3.707	3.745
SAT Math Score	514	614.026	614.481
SAT Verbal Score	500	567.464	562.124
Entering Units	---	37.326	41.426

Note. Author derived national demographic data from the IPEDS Data Center 2008 data collection; national sample includes public degree-granting not-for-profit institutions and full-time undergraduate totals. SAT information retrieved from: https://reports.collegeboard.org/pdf/total-group-2016.pdf

Table 3.2
Instructor Descriptive Statistics

Instructor-Level	**All Courses** (1) Teaching Faculty	(2) Research Faculty	(3) Lecturers	**Courses in the Analytical Sample** (4) Teaching Faculty	(5) Research Faculty	(6) Lecturers
Instructor is Female	0.500	0.349	0.506	0.583	0.346	0.538
White	0.775	0.681	0.651	0.812	0.706	0.679
Black	0.037	0.027	0.018	0.021	0.034	0.023
Hispanic	0.075	0.064	0.072	0.062	0.061	0.057
Asian	0.087	0.201	0.177	0.104	0.179	0.167
Other	0.025	0.027	0.082	0.000	0.020	0.075
Instructor Age in 2008	43.150	45.011	38.914	45.208	46.863	41.894
Total Years at University (maximum)	15.375	15.765	6.275	17.250	17.792	8.274
Full-Time	0.750	0.916	0.110	0.729	0.929	0.120
Monthly Pay Rate	7021.359	10390.699	5292.696	7105.434	9378.239	5005.450
Average number of courses per term	2.318	1.463	1.477	2.418	1.613	1.693
Average number of course credits per term	8.204	5.503	5.595	8.644	6.112	6.621
N	80	1,329	1,090	48	592	442

Note. Data are on instructors in the full sample who are tied to a course taken by 2008-2013 cohorts. Full-time employed defined as worked as full-time instructor during at least half of the terms employed in the institution.

Table 3.3
Summary Statistics for Outcome Measures by Faculty Type, Academic Year 2008-2016

	Teaching Faculty	Research Faculty	Lecturer
Panel A. Contemporaneous Course Outcomes			
Course grade (0 to 4 grading scale)	2.712	2.824	2.948
Panel B. Subsequent Enrollment Outcomes			
Enrolled in subsequent course in the same field	0.850	0.839	0.804
Average grade of subsequent course	2.720	2.807	2.924
Panel C. Subsequent Class Outcomes			
Course grade (0 to 4 grading scale)	2.695	2.798	2.916
Total credits earned in the same field	24.76	26.25	17.74
Declared a major in the same field	0.200	0.299	0.165

Note: The contemporaneous course sample is restricted to the first college-level course taken by each student in each field of study. We exclude courses with pass/fail grades.

Table 3.4
Probability of Taking an Introductory Course with Different Types of Instructors

	(1) Research Faculty	(2) Teaching Faculty
Student-level characteristics:		
Female	-0.008*	0.002
	(0.004)	(0.002)
Black	-0.003	0.005
	(0.006)	(0.008)
Hispanic	0.004	0.002
	(0.004)	(0.007)
Asian	-0.003	-0.006
	(0.004)	(0.004)
Other	-0.009	-0.018
	(0.010)	(0.015)
Age when taking the course	0.000	0.001
	(0.001)	(0.001)
Low-income status	0.006**	-0.004*
	(0.003)	(0.002)
First-gen status	0.001	0.004
	(0.001)	(0.003)
SAT Math Score	0.000	0.000
	(0.000)	(0.000)
SAT Verbal Score	-0.000	0.000
	(0.000)	(0.000)
HS GPA	0.004	-0.010
	(0.004)	(0.010)
Entering Units	0.002	0.002**
	(0.001)	(0.001)
Course-section level characteristics:		
Average enrollment	0.001***	0.001**
	(0.000)	(0.000)
Lower-division course	-0.268***	-0.094
	(0.089)	(0.067)
Course is online	-0.047	-0.411**
	(0.247)	(0.175)
Course is offered in the morning	0.027	-0.115
	(0.072)	(0.090)

Note. All regressions control for high school fixed effects, course fixed effects, term fixed effects, and cohort fixed effects. We also include section level controls. The base group for regressions is lecturer. The base group for race is white. Standard errors are clustered at the subject level. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.5
Contemporaneous Course Outcomes

	(1) Grade (incl. W)	(2) Grade (excl. W)
Research Faculty	-0.015	-0.013
	(0.038)	(0.038)
Teaching Faculty	-0.060	-0.062
	(0.036)	(0.036)
Sample Mean	2.846	2.855
N	80980	80732
F-test	p < 0.001	p < 0.001
Term FEs	X	X
Introductory Course FEs	X	X
Student FEs	X	X

Note. Base group is lecturer. Controls for all models include course section characteristics of the introductory course including enrollment size, delivery method, other students' average SAT score in the class, and day and time of class. We also include a control for whether the course is within a student's intended major. Classes on pass fail grading system are excluded. Standard errors are two-way clustered at the student and subject level. Robust standard errors in parentheses. *** p<.01, ** p<0.05, * p<.10.

Table 3.6
Impact of Different Types of Instructors on Course Performance: Grade Distribution

Panel A. Contemporaneous Course Performance

Outcome	(1) Grade: A- or above	(2) Grade: B- or better	(3) Grade: C- or better	(4) Grade: D- or better	(5) Grade-numeric
Research Faculty	-0.008	0.000	-0.002	0.006	-0.013
	(0.017)	(0.015)	(0.008)	(0.007)	(0.038)
Teaching Faculty	-0.002	-0.017	-0.008	-0.027**	-0.062
	(0.015)	(0.017)	(0.007)	(0.007)	(0.036)
N	80732	80732	80732	80732	80732
Term FEs	X	X	X	X	X
Introductory Course FEs	X	X	X	X	X
Student FEs	X	X	X	X	X

Panel B. Subsequent Course Performance

	(1)	(2)	(3)	(4)	(5)
Research Faculty	0.002	0.005	0.008	0.002	0.021
	(0.008)	(0.013)	(0.007)	(0.006)	(0.020)
Teaching Faculty	0.015	0.018	0.018*	0.003	0.053*
	(0.012)	(0.015)	(0.009)	(0.006)	(0.030)
N	66836	66836	66836	66836	66836
Term FEs	X	X	X	X	X
Introductory Course FEs	X	X	X	X	X
Student FEs	X	X	X	X	X
Next-class-taken FEs	X	X	X	X	X

Note. All models include students' first course in a field of study during their first term. Base group is lecturer. Controls for all models include course section characteristics of the introductory course including enrollment size, delivery method, other students' average SAT score in the class, and day and time of class. We also include a control for whether the course is within a student’s intended major. Students receiving a "W" in the introductory course are dropped from the sample. Students receiving a "W" in the next course are dropped from the sample. Classes on pass fail grading system are excluded. Standard errors are two-way clustered at the student and subject level. Robust standard errors in parentheses. *** p<.01, ** p<0.05, * p<.10.

Table 3.7
Subsequent Outcomes

	(1) Enrolled in Subsequent Course	(2) Average Grade of Second Course	(3) Subsequent Course Grade	(4) Total Credits Earned in a Field of Study	(5) Declared a Major in the Same Field of Study
Research Faculty	-0.004	0.007	0.021	0.911	-0.001
	(0.008)	(0.007)	(0.020)	(0.588)	(0.007)
Teaching Faculty	0.012	-0.002	0.053*	1.250*	0.002
	(0.013)	(0.008)	(0.030)	(0.696)	(0.005)
Sample Mean	0.828	2.837	2.827	23.09	0.241
N	80980	67056	66836	80980	78980
F-test	p < 0.001	p < 0.001	p < 0.001	p < 0.001	p < 0.001
Term FEs	X	X	X	X	X
Introductory Course FEs	X	X	X	X	X
Student FEs	X	X	X	X	X
Next-class-taken FEs			X		

Note. Base group is lecturer. Controls for all models include course section characteristics of the introductory course including enrollment size, delivery method, other students' average SAT score in the class, day and time of class. We also include a control for whether the course is within a student's intended major. Students receiving a "W" in the introductory course are dropped from the sample. Classes on pass fail grading system are excluded. Standard errors are two-way clustered at the student and subject level. Robust standard errors in parentheses. *** p<.01, ** p<0.05, * p<.10.

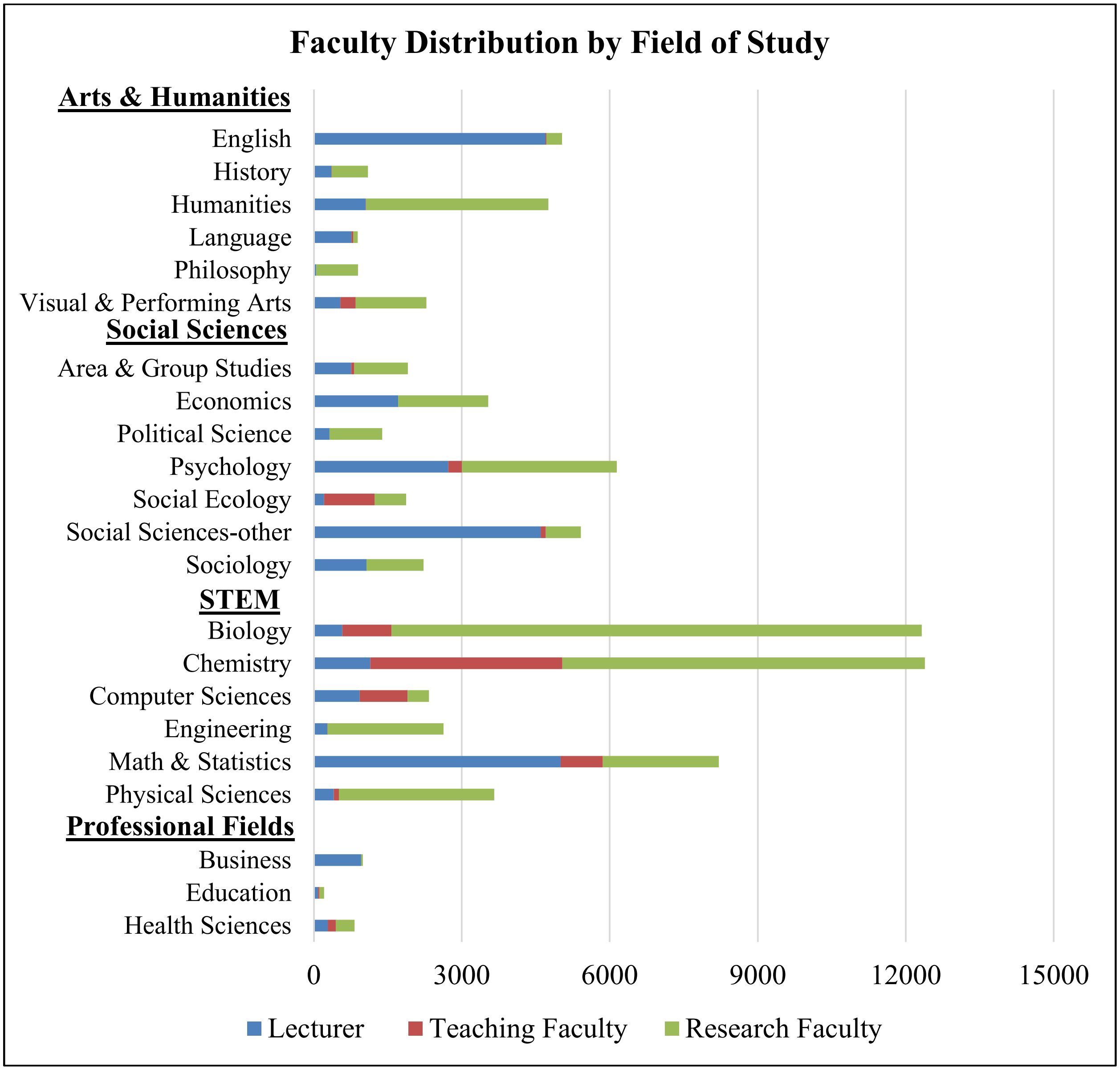

Figure 3.1 Faculty distribution by field of study.

Appendix

Appendix Table 3.1
Faculty Titles by Faculty Type

Teaching Faculty	Research Faculty	Lecturers
Emeritus Lecturer with Security of Employment	Emeritus Professor	Continuing Appointment Temp August (Unit-18)
Senior Lecturer with Security of Employment	Acting Assistant Professor	Supervisor of Teacher Education (Unit-18)
Emeritus Senior Lecturer with Security of Employment	Research Professor	Summer Session Lecturer (Unit-18)
Lecturer with Security of Employment	Acting Professor	Lecturer (Unit-18)
Senior Lecturer with Potential Security of Employment	Assistant Professor	Senior Lecturer (Unit-18)
Lecturer with Potential Security of Employment	Professor	
Lecturer with Potential Security of Employment (Unit-18)	Associate Professor	
	Acting Associate Professor	

Appendix Table 3.2
Faculty Distribution by Field of Study

Field of Study (Subject)	Teaching Faculty	Research Faculty	Lecturer
Area & Group Studies	65	1084	757
Biology	1001	10747	577
Business	0	38	958
Chemistry	3894	7345	1146
Computer Science	969	433	930
Economics	0	1827	1711
Education	26	94	89
Engineering	4	2350	276
English	18	314	4699
Health	161	379	286
History	5	733	358
Humanities	4	3698	1052
Language	33	92	764
Math & Statistics	856	2347	5005
Philosophy	0	851	45
Physical Sciences	104	3149	406
Political Science	0	1068	318
Psychology	280	3134	2727
Social Ecology	1022	642	208
Social Sciences	106	709	4600
Sociology	0	1153	1071
Visual & Performing Arts	313	1432	538

Each cell indicates the total number of course enrollments for each faculty type within each subject. Includes the analytical sample.

Appendix Table 3.3
Summary Statistics: Analytical Sample and Restricted Sample

	AUCC-Analytical Sample	AUCC-Restricted Sample
Student demographic characteristics:		
Female	0.549	0.557
White	0.203	0.181
Black	0.027	0.025
Hispanic	0.206	0.203
Asian	0.553	0.583
Other	0.011	0.009
Student Age	19.289	18.734
Low-income Status	0.317	0.323
First-generation Status	0.450	0.452
HS GPA	3.745	3.811
SAT Math Score	614.479	616.717
SAT Verbal Score	562.122	562.822
Entering Units	41.416	31.429

Note. Restricted sample includes observations that contribute to the fixed effects estimation.

Appendix Table 3.4
Impact of Different Types of Instructors on Total Credits

	(1) Earned at least 1 credit	(2) Earned 8 credits or more	(3) Earned 12 credits or more	(4) Total Credits-continuous
Research Faculty	-0.005	0.006	0.019*	0.911
	(0.008)	(0.010)	(0.011)	(0.588)
Teaching Faculty	0.013	0.044**	0.038**	1.250*
	(0.014)	(0.017)	(0.017)	(0.696)
N	80980	80980	80980	80980
Term FEs	X	X	X	X
Introductory Course FEs	X	X	X	X
Student FEs	X	X	X	X

Note. Earning at least 1 credit does not include the introductory course. Base group is lecturer. Controls for all models include course section characteristics of the introductory course including enrollment size, delivery method, other students' average SAT score in the class, and day and time of class. We also include a control for whether the course is within a student’s intended major. Classes on pass fail grading system are excluded. Standard errors are two-way clustered at the student and subject level. Robust standard errors in parentheses. *** p<.01, ** p<0.05, * p<.10.

Appendix Table 3.5
Relevant Literature

	Xu & Solanki	Xu & Solanki	Ran & Xu (in press)	Figlio et al. (2015)	Hoffman & Oreopoulos (2009)
Setting	4-year public institution	4-year public institution	4-year public institution	4-year private institution	4-year public institution
	Lecturer v. Teaching faculty	**Lecturer v. Tenure-track/Tenured Research Faculty**	Long-term contingent faculty v. Tenure Track Faculty	Lecturer v. Tenure-track/Tenured Faculty	Lecturer v. Full Professor
<u>Outcomes</u>					
Current course grade	0.062	0.015	0.170***	---	-1.171
	(0.036)	(0.038)	(0.016)	---	(0.515)
	M = 2.855	M = 2.855	M = 2.47	---	M = 69.18
Enrolled in subsequent course	-0.012	0.005	-0.010**	0.073***	---
	(0.013)	(0.008)	(0.004)	(0.024)	---
	M = 0.828	M = 0.828	M = 0.43	M = 0.74	---
Subsequent course grade	-0.053*	-0.021	-0.016	0.060***	---
	(0.030)	(0.020)	(0.010)	(0.008)	---
	M = 2.827	M = 2.827	M = 2.48	M = 3.39	---

Note. Point estimates reflect lecturer effect.

CHAPTER 4

Summary and Conclusions

Overall Summary

The three studies outlined above sought to better understand institutions (specifically, four-year public research institutions), the students they serve, and the efforts they take to improve student outcomes. Specifically, the three studies focus on evaluating institutional policies and programs intended to support student success. I found that interventions based on theories of student motivation are linked to academic achievement in college and that impacts are more pronounced for students most at risk for poor academic achievement. However, I also found that these interventions may not, in their current form, be best suited to help students at broad-access institutions.

I also found that holistic student support programs such as EASE improved *STEM* students' experience in college—and, more specifically, in their major—by fostering feelings of belonging. Sense of belonging is a likely mechanism for increases in academic achievement among EASE students, and it therefore supports the notion that sense of belonging is a critical dimension of success at college. Lastly, I found that student learning isn't correlated with faculty type; students learn equally well from tenure-track teaching faculty, research faculty, and lecturers. Institutional policies that aim to reallocate faculty responsibilities could therefore be a promising approach for maintaining excellence in regard to both research and teaching; they could also be a means of facilitating better research-teaching balance for those employed.

The three studies in my dissertation are distinct but, as a whole, inform research literature about higher education policy more broadly because they focus on issues of significance to institutional administrators and practitioners. The studies detailed in this dissertation, in addition to their findings, also open the door to promising future questions and lines of research, which I outline next.

Future Directions

Study 1. A significant contribution of the social-psychological intervention literature is that it has brought heightened attention to the need to focus on social-psychological issues in the college context. A good approach moving forward may be to determine how best to alter the educational environment to better embody principles and practices of social-psychological theories. It may be particularly beneficial to test how student learning and engagement change when instructors leverage mindsets and motivation in their classroom and instructional practices.

For example, the University of California system recently implemented the First Generation College Initiative, a program meant to generate awareness of first-generation students and faculty on campus. The program has motivated faculty to make their first-generation status identifiable to students, encouraging faculty to wear t-shirts with the graphic *#Firstgen College Grad Faculty* in class, for example. Future research could investigate whether subtle messages like this foster feelings of belonging among first-generation college students and whether this initiative is related to student engagement and achievement.

As another example, future research could also examine whether specific instructional strategies—such as grading policies—encourage a growth mindset. This topic has been explored in K-12 literature, with studies focusing on comparing traditional grading practices (i.e., those in which students focus on earning grades) and standard grading practices (i.e., those in which students focus on learning an objective). Less empirical work has been done in college classrooms, but future studies could involve instructors experimenting with different grading schemes, measuring whether they impact students' framing of challenges.

Study 2. In light of the findings discussed in chapter 2, a future study about learning communities and their impact on student outcomes across different STEM disciplines could be both interesting and informative. Other STEM fields such as engineering, in which females and URM students are less represented than they are in the biological sciences, could also provide a promising context for exploring treatment effect heterogeneity.

Another future extension of this research could focus on identifying the mechanisms that make a learning community work. Numerous studies have documented the positive effects of social networks on academic outcomes (Sacerdote, 2001; Stadtfeld et al., 2018) specifically related to academic achievement and attainment (Carbonaro & Workman, 2016; Vaquera & Kato, 2008). Advances in social network analysis could test the social network mechanisms (development of friendships and academic relations, for example) at the heart of learning communities such as EASE, examining how these networks relate to academic outcomes. A study with components like this can inform program design for future learning communities.

Lastly, another mechanism worth exploring is that of EASE peer mentors. Future work could follow what has been done by previous scholars, gathering qualitative data from EASE participants in an attempt to better understand aspects of the mentor/mentee relationship that are particularly effective. Also, a particular line of research shows that having a same-gender or same-race instructor improves student learning and engagement

(Carrell, Page, & West, 2010; Dee, 2005; Fairlie, Hoffman, & Oreopoulos, 2014; Hoffman & Oreopoulos, 2009; Solanki & Xu, 2018). Mentor/mentee relationships is not a highly researched area; thus, research in this sector could be informative for future learning community design.

Study 3. Study 3 focuses on evaluating faculty, as opposed to evaluating policy change responsible for the tenure-track teaching faculty category in the UC system. However, the policy change and its impacts are equally interesting. A future study—depending on data availability—could test the efficacy of this policy change and examine whether the implementation of tenure opportunities in the UC system has changed instructor effectiveness and the quality of incoming faculty.

Results from study 2 also motivate potential future research about contingent faculty. Today, current evidence about contingent faculty and their impact on undergraduate education is fairly mixed. In fact, I found that tenure-track teaching faculty and lecturers look quite similar in terms of student outcomes, which is a unique finding in the field. The mixed results of studies about faculty impact on student outcomes suggest quality differences in the contingent faculty pool and also indicate that certain institutions have a greater ability to attract better candidates. One way to test this hypothesis is to examine differences among contingent faculty in terms of their background, experience, and credentials, exploring whether these differences are related to student outcomes. For example, at a number of teaching institutions such as those in the California State University system, more than half of faculty hired are contingent, and they vary in terms of educational background and highest degree attained. Future research could use this type of data to examine quality differences among contingent faculty, and a study like this could certainly inform other institutions about contingent faculty hiring policies.